

137302

ANKET ARAŐTIRMALARINDA YANITLAMAMA VE
ÇİFT ÖRNEKLEME YÖNTEMİ

Aylin ALKAYA

YÜKSEK LİSANS TEZİ
(İSTATİSTİK)

GAZİ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

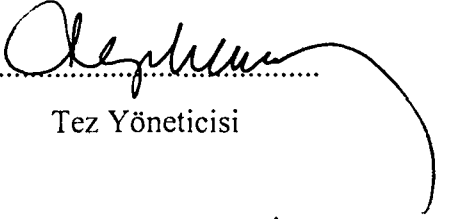
137302

T.C. YÜKSEKÖĞRETİM KURULU
DOKÜMANTASYON MERKEZİ

Temmuz 2003

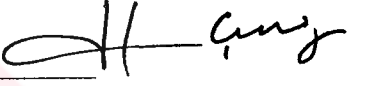
ANKARA

Aylin ALKAYA tarafından hazırlanan ANKET ARAŞTIRMALARINDA YANITLAMAMA VE ÇİFT ÖRNEKLEME YÖNTEMİ adlı bu tezin Yüksek Lisans tezi olarak uygun olduğunu onaylarım.

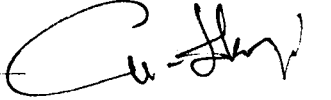

Tez Yöneticisi

Prof. Dr. A. Alptekin ESİN

Bu çalışma, jürimiz tarafından İSTATİSTİK Anabilim Dalında YÜKSEK LİSANS tezi olarak kabul edilmiştir.

Başkan: : PROF. DR. HÜLYA CINGİ 

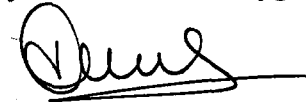
Üye : PROF. DR. ALPTEKİN ESİN 

Üye : PROF. DR. MÜSLİM EKİNİ 

Üye : _____

Üye : _____

Bu tez, Gazi Üniversitesi Fen Bilimleri Enstitüsü tez yazım kurallarına uygundur.



İÇİNDEKİLER

	Sayfa
ÖZET	i
ABSTRACT	iii
TEŞEKKÜR	v
ÇİZELGELERİN LİSTESİ	vi
ŞEKİLLERİN LİSTESİ	vii
EKLERİN LİSTESİ	viii
1.GİRİŞ	1
2. ANKET ARAŞTIRMALARINDA YANITLAMAMA.....	4
2.1. Anket Yöntemi.....	5
2.1.1. Araştırma sorununun teşhisi ve tanımı	6
2.1.2. Örneklem tasarımı/ tamsayım	7
2.1.3. Veri toplama yöntemi seçimi.....	8
2.1.4. Anket formunun hazırlanması.....	14
2.1.5. Araştırmanın planlanması ve organizasyonu	16
2.1.6. Alan uygulaması.....	17
2.1.7. Verilerin işlenmesi ve analizi.....	19
2.1.8. Araştırma raporunun sunumu	20
2.2. Araştırmalarda Yanıtlamama.....	20
2.2.1. Yanıtlamamanın belirlenmesi	21
2.2.2. Yanıtlamama türleri.....	23
2.2.3. Yanıtlamama kaynakları.....	24
2.2.4. Yanlar ve örneklem dışı hatalar.....	27

2.2.5. Yanıtlamamanın etkileri	31
3. YANITLAMAMA İÇİN ÇİFT ÖRNEKLEME YÖNTEMİ	36
3.1. Hansen ve Hurwitz Yöntemi	37
3.1.1. Yiğın ortalamasının tahmini ve tahminin varyansı	40
3.1.3. n ve k 'nın optimum değerleri	45
3.1.3. Oran ve regresyon tahmini.....	49
3.2. Cochran'ın Yöntemi.....	51
3.2.1. Yiğın ortalamasının tahmini ve tahminin varyansı	51
3.2.2. n , n_2' ve k 'nın optimum değerleri	53
3.2.3. W_2 yanıtlamayanlar oranı bilinmiyorsa.....	56
3.2.4. Oran ve regresyon tahmin edicileri	56
3.3. Srinath'ın Yöntemi.....	57
3.3.1. Yiğın ortalamasının tahmini ve tahminin varyansı	58
3.3.2. n , n_2' ve k 'nın optimum değerleri	61
3.3.3. $L > 1$ için altörnekleme kuralı	63
3.4. Deming'in Yöntemi	66
4. UYGULAMA.....	72
4.1. Birinci Soru Grubu.....	74
4.1.1. n örnek çapının belirlenmesi.....	74
4.1.2. Tek örnekleme yöntemi.....	76
4.1.3. Yanıtlamama için çift örnekleme yöntemi	79
4.2. İkinci Soru Grubu	103
4.2.1. n örnek çapının belirlenmesi.....	103
4.2.2. Tek örnekleme yöntemi.....	104

4.2.3. Yanıtlamama için çift örnekleme yöntemi	106
6. SONUÇ VE ÖNERİLER.....	126
KAYNAKLAR.....	132
EKLER.....	136
ÖZGEÇMİŞ	167



ANKET ARAŐTIRMALARINDA YANITLAMAMA VE ÇİFT ÖRNEKLEME YÖNTEMİ

(Yüksek Lisans Tezi)

Aylin ALKAYA

GAZİ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

Mayıs 2003

ÖZET

Anket arařtırmalarında yanıtlanmadan dolayı oluşan kayıp verinin iki temel tür ayırımı yapılmıřtır. Birinci türü, birim yanıtlanmamadır; örneđi oluřturan bazı birimlere iliřkin hiçbir veri yoktur. İkinci türü ise, soru yanıtlanmamadır; örneđi oluřturan birim veya birimlere iliřkin bir veya birden fazla soru veya deđiřken verisi yoktur. Bu çalıřmada, birim yanıtlanmama ve kaynakları, nedenleri, nelere yol açtıđı incelenmeye çalıřılmıřtır. Yanıtlanmama tahminlerin kalitesini ve güvenilirliđini azaltmaktadır. Yalnızca yanıtlayanlara dayalı yapılan parametre tahminleri, yanıtlanmayanların yanıtlayanlardan farklı olması nedeniyle, yanlı olur. Arařtırmalarda yanıtlanmayan birimlerin varlıđı, örnek çapının küçülmesine ve arařtırma maliyetinin artmasına da neden olur. Bu nedenle, yanıtlanmamanın ortaya çıkmaması için çalıřmalar yürütülmeli ve varlıđında tahminler üzerindeki etkilerinin minimum yapılması için düzeltmeler yapılmalıdır. Bu çalıřmada, yanıtlanmama yanlılıđının giderilmesi ve yanıt oranının arttırılması için önerilen yöntemlerden biri olan çift örnekleme yöntemi pek çok arařtırmaya ışık tutacak şekilde ele alınmıřtır. Tezde, yanıtlanmama için çift örnekleme yönteminde; yanıtlayanlar ve yanıtlanmayanlar iki ayrı tabakaya sınıflandırılmıř ve yanıtlanmayanlar tabakasından bir tesadüfi altörnek seçilmifitir. Seçilen altörnekten, ilk örnektekine kıyasla, yanıtlanmayan birimlerinden yanıt almak için yoğun çabanın harcandıđı bir veri toplama

yönteminin yanında, daha farklı yöntem veya yöntemler kullanılarak yanıt alınmaya çalışılır.



Bilim Kodu :
Anahtar Kelimeler : yanıtlanmama, birim yanıtlanmama, çift örnekleme
Sayfa Adedi : 165
Tez Yöneticisi : Prof. Dr. A. Alptekin ESİN

**NONRESPONSE IN SURVEY RESEARCH AND
DOUBLE SAMPLING TECHNIQUE**

(M.Sc. Thesis)

Aylin ALKAYA

**GAZI UNIVERSITY
INSTITUTE OF SCIENCE AND TECHNOLOGY**

May 2003

ABSTRACT

In survey researches a distinction between two major types of missing data in surveys due to nonresponse had been made. First type is, unit nonresponse; for some sample units no entire data is available. The second type is, item nonresponse; sample unit or units have missing data for one or more questions or variables. In this study, unit nonresponse and its sources, reasons, what it caused was tried to be investigated. Nonresponse reduce the quality and reliability of the estimates. The parameter estimates that only depend on respondents will be biased because of the difference between nonrespondents and respondents. In researches the existence of nonrespondents, cause a decrease in sample size and an increase in research costs. For this reason, to prevent nonresponse occurrence efforts must be carried out and in its presence, to minimize its effects on estimations, corrections must be done. In this study, the double sampling method, which is one of the suggested methods for removing nonresponse bias and increasing response ratio, was evaluated to guide for many researches. In thesis, in double sampling method for nonresponse, the response and nonresponse units were classified into different clusters and one random subsample was taken from nonresponse cluster. From the taken subsample, compared to first sample, besides a data collection method

that more efforts are spend to get answers from nonresponse units, some different methods are used to get answers.



Science Code :

Key Words : nonresponse, unit nonresponse, double sampling

Page Number: 165

Adviser : Professor Dr. A. Alptekin ESİN

TEŐEKKÜR

Çalıřmalarım süresince deęerli yardım ve katkılarıyla beni yönlendiren, kıymetli zamanımı bana ayırıp titizlikle tezi inceleyen tez danışmanım, hocam Prof.Dr.A.Alptekin ESİN'e, yine kıymetli fikirlerinden yararlandığım deęerli hocam Yrd.Doç.Dr.Celal AYDIN'a teőekkürü bir borç bilirim. Anket sorularının hazırlanmasında önerilerde bulunan E.Ü. İ.İ.B.F. öğretim üyesi Yrd.Doç.Dr.Mahmut ÖZDEVECİOĞLU'na, anketin uygulanma olanağını sağlayan Nevşehir Tekel Fabrikası çalışanlarına ve anketi yanıtlama nezaketi gösteren tüm fabrika personeline teőekkür ederim. Ayrıca bu çalışmam esnasında yardım ve desteklerini esirgemeyen arkadaşlarım Fatih AYDOĞAN'a, E.Ü. Nevşehir İ.İ.B.F. öğretim üyesi Yrd.Doç.Dr.Suzan ÇOBAN'a, S.Zulal ÖZEREN'e ve Nalan ÇAĞIR'a ve manevi destekleriyle beni hiçbir zaman yalnız bırakmayan aileme teőekkür ederim.

ÇİZELGELERİN LİSTESİ

Çizelge	Sayfa
Çizelge 3. 1. Beklenen maliyetlerin kıyaslanması	66
Çizelge 3. 2. Deming modeli için göstermelik parametre değerleri.....	69
Çizelge 3. 3. İlk girişimden sonra yanıtlamayanların optimal 0,60 çift örnekleme varsayımıyla araştırma sonuçları.....	71
Çizelge 4. 1. Öntest sonucunda birinci soru grubundan elde edilen istatistik değerleri	75
Çizelge 4. 2. Yanıtlayanlar tabakasına dayalı hesaplanan istatistik değerleri	77
Çizelge 4. 3. Altörnek birimlerine ilişkin istatistik değerleri.....	84
Çizelge 4. 4. Posta anketini yanıtlayanlar ve yanıtlamayanlar arasından seçilen altörnek birimlerine ilişkin istatistik değerleri	84
Çizelge 4. 5. Posta anketini yanıtlayanlara ve yanıtlamayanlar arasından seçilen $n_{25}=30$ ön örnek birimlerine ilişkin istatistik değerleri.....	94
Çizelge 4. 6. Altörnek birimlerine ilişkin istatistik değerleri.....	97
Çizelge 4. 7. Posta anketini yanıtlayanlar ve yanıtlamayanlar arasından seçilen altörnek birimlerine ilişkin istatistik değerleri	97
Çizelge 4. 8. Birinci soru grubu sonuç istatistikleri	101
Çizelge 4. 9. Öntest sonucunda elde edilen istatistik değerleri.....	103
Çizelge 4. 10. Yanıtlayanlar tabakasına dayalı hesaplanan istatistik değerleri.....	104
Çizelge 4. 11. Altörnek birimlerine ilişkin istatistik değerleri.....	108
Çizelge 4. 12. Posta anketini yanıtlayanlar ve yanıtlamayanlar arasından seçilen altörnek birimlerine ilişkin istatistik değerleri	109
Çizelge 4. 13. Posta anketini yanıtlayanlar ve yanıtlamayanlara ilişkin istatistik değerleri	111
Çizelge 4. 14. Posta anketini yanıtlayanlara ve yanıtlamayanlar arasından seçilen altörnek birimlerine ilişkin istatistik değerleri	113

Çizelge 4. 15. Posta anketini yanıtlayanlar ve yanıtlayanlar arasından seçilen $n_{20}=30$ ön örnek birimlerine ilişkin istatistik değerleri.....	116
Çizelge 4. 16. Altörnek birimlerine ilişkin istatistik değerleri.....	118
Çizelge 4. 17. Posta anketini yanıtlayanlar ve yanıtlayanlar arasından seçilen altörnek birimlerine ilişkin istatistik değerleri	118
Çizelge 4. 18. İkinci soru grubu sonuç istatistikleri	123



ŞEKİLLERİN LİSTESİ

Şekil	Sayfa
Şekil 2. 1. Tam yanıt olması durumu.....	21
Şekil 2. 2. Yanıtlamama durumu	22
Şekil 2. 3. Yanıtlamama durumu	22
Şekil 2. 4. Yanıtlamama durumu	22
Şekil 3. 1. Yanıtlamama olması durumunda çift örnekleme yöntemi	39



EKLERİN LİSTESİ

Ek	Sayfa
EK-1. Anket Formu	137
EK-2. http://www.Bls.Census.Gov/Sipp/Qp/C5.Htm İnternet Kaynağı.....	142
EK-3. http://wwwlib.umi.com/dissertations/preview_page/NQ47726/13 İnternet Kaynağı	160



1.GİRİŞ

Anket arařtırmalarının amacı, arařtırmanın uygulandıđı birimlerden yanıt alınması olduđuna göre yanıtlanmama istenilmeyen bir durumdur ve yanıtlanmama örnekleme dıřı bir hata olarak kabul edilir (1).

Yanıtlanmama olması veri kümesinde eksikliklerin olması demektir. Eksik veri olması standart tam veri yöntemlerinin kullanılamayacađı anlamına gelir. Yanıtlanmama olması yana neden olur çünkü genellikle yanıtlanmayanlar yanıtlayanlardan farklıdır ve yanıtlanmamadan kaynaklanan yanı ihmal etmek zordur (2). Yanıtlanmayanların yanıtlayanlardan farklı olması nedeniyle yanıtlayanlar örneğinin yığını temsil etmesi gücü düşüktür (3).

Bu çalışmada yanıtlanmama olması durumunda parametre tahminlerinin çift örnekleme yöntemiyle hesaplanması amaçlanmıştır.

Ülkemizde yanıtlanmama üzerine oldukça az sayıda çalışma yapılmıştır. Yanıtlanmama sorununa ilk HIPS (1980) çalışmasında yer verilmiştir (4). Türkiye’de yanıtlanmama ile ilgili yapılan çalışmaların neler olduđu ise çizelge halinde Survey Nonresponse Models And Applications In Turkey çalışmasında verilmiştir (4).

Yanıtlanmama, arařtırma kapsamında bulunan bir kiři veya birim anketi yanıtlanmıyorsa (birim yanıtlanmama (unit nonresponse)) veya kiři veya birim(ler) anket çalışmasına katılıyor ancak en azından bir deđiřkeni veya soruyu yanıtlanmıyorsa (soru yanıtlanmama (item nonresponse)) ortaya çıkar (5). Çalışma kapsamında yanıtlanmama türlerinden birim yanıtlanmama incelenmiştir.

Arařtırma yöntemleriyle uğrařanlar yanıtlanmamaya iliřkin güçlükleri gidermek için çeřitli yöntemler geliřtirmişlerdir. Bu yöntemler arasında en kayda deđerleri, deđerik zamanlarda birimlerle görüşme yapılabilmesi olasılıklarını da içeren yanıt modelleri ve yanıtlanmayanlardan altörnek alınmasına dayanan çift örnekleme yöntemidir. Ayrıca, yanıtlanmayanların yanıtlayanlara benzediđi varsayımı esas alınıp bazı

düzeltilme yöntemleriyle yanıt oranını arttıran tekniklerinde geniş kullanım alanı bulunduğu görülmektedir (6,7,8). Bu çalışmada yanıtlamama için çift örnekleme yöntemi incelenmiştir.

Tezin birinci bölümünde anket araştırmasının süreçleri ve yanıtlamamanın ne olduğu verilmeye çalışılmıştır. Yanıtlamama yalnızca görüşme anında yanıt alınmaması anlamında olmayıp, görüşme öncesinde ve sonrasındaki bütün aşamaları kapsayarak analizlerde kullanılacak son bilgi kütüğüne ulaşabilen yanıtları da içerir (8). Bu nedenle birinci bölümde, anket araştırmasının geçirdiği aşamalar, yanıtlamamanın minimum, veri kalitesinin ise maksimum olması sağlanacak şekilde düzenlenmesi ve yürütülmesi gerekliliğine değinilerek verilmeye çalışılmıştır. Bu bölümün ikinci kısmında ise; yanıtlamama ile ilgili temel tanımlar, yanıtlamama kaynakları ve hatalar içindeki yerinin ne olduğu üzerinde durulmuştur.

Çalışmanın ikinci bölümünde, yanıtlamama yanlılığının giderilmesi ve yanıt oranını artırılması için önerilen yöntemlerden Hansen ve Hurwitz (9), Cochran (10), Srinath (11) ve Deming (12)'in çift örnekleme yöntemi incelenmiştir.

Üçüncü bölümde ise bu çalışmanın uygulama kısmı yer almaktadır. Bu bölümde yanıtlamama için çift örnekleme yönteminin bir uygulamasının yapılması amaçlanmıştır. Nevşehir Tekel Fabrikası çalışanlarına iş ortamı profilinin tespit edilmesi amacıyla posta anketi yöntemi uygulanmıştır. Posta anketi diğer anket yöntemlerine kıyasla en düşük maliyetli veri toplama yöntemi olmasına karşın, en düşük yanıt oranına sahip bir anket yöntemidir. Uygulamada posta anketi yönteminin tercih edilmesinin nedeni, posta anketlerine olan yanıt düzeyinin ne olacağının tespit edilmesi ve yanıtlamama olması durumunda çift örnekleme yönteminin bir uygulamasının yapılabilmesidir.

Son bölümde uygulama kısmına ait sonuç ve değerlendirmeye ilişkin bilgiler yer almaktadır. Bu bölümde, yapılan uygulama çalışması sonucunda yanıtlamama olması durumunda yanıtlamayanlardan veri toplamak üzere girişimlerin yapılması gerektiği,

bunun için önerilen çift örnekleme yönteminin kullanılacağı ve bu yöntemle yapılan tahminlerin yansız tahminler olduğu vurgulanmıştır.



2. ANKET ARAŞTIRMALARINDA YANITLAMAMA

Araştırma yapmaya gerek duymanın değişik nedenleri vardır. Bunların başında ilgi ve zorunluluk gelir. Araştırma, bir soruna güvenilir çözümler getirecek yansız, tutarlı ve güvenilir tahminlerde bulunmak amacıyla verilerin planlı olarak toplanması, sınıflandırılması ve değerlendirilmesi sürecidir (13).

Anketlerin amacı, toplumdaki birimlerin (ya da bireylerin) tümü ya da bir bölümü ele alınarak, bu birimlerin bir ya da birden fazla özelliğini (değişkenini) ölçerek, tartarak ya da sayarak sayısal değerlerini saptamak ve bu sayısal değerleri kullanarak belirli sonuçlara ve kararlara ulaşmaktır.

Araştırmaların amacı elde edilen verilerden gerçek değeri elde etmektir. Fakat, yanıtlayıcıdan, anketörden, yanıtlayıcı-anketör etkileşiminden ve anket formundan kaynaklanan hatalar nedeniyle, araştırmacı gerçek değer yerine kaydedilen değeri elde edebilmektedir (14). Gerçek değer ile kaydedilen değer arasındaki fark, bir araştırmanın örnekleme dışı hatasını verir.

Groves (1), hataların azaltılmasının; istenen bilginin yüksek kalitede alınmasını sağlayacak soruların hazırlanmasıyla, yığını en iyi temsil edecek örnekleme yönteminin seçilmesiyle, yanıtlayıcılar üzerinde anketör etkisinin azaltılmasıyla sağlanabileceğini belirtmiştir. Ayrıca örneğe seçilen tüm birimlerden tam yanıt alınmasının, tahminlerin güvenilirliğinde ölçüt olacağını vurgulamıştır.

Anket yöntemlerinin amacı, araştırmanın uygulandığı birimlerden yanıt alınması olduğuna göre, yanıtlamama olması istenilmeyen bir durumdur (1). Dolayısıyla, örnekleme dışı bir hata olan yanıtlamamanın giderilmesi veya minimum yapılması için gerekli düzenlemeler yapılmalıdır.

Bu nedenle örnekleme dışı hataların kontrolünde yapılması gereken bazı temel işlemler söz konusudur. Bu noktalar aşağıda verilmiştir.

1. Görüşmecilerin seçimi, eğitimi ve alanda kontrol edilmesi,
2. Görüşme sürecinin kontrolü,
3. Görüşme saatlerinin seçimi, (Gündüz saatleri çalışanları ya da öğrencileri evlerinde bulmada seçilmiş kötü saatlerdir. En uygun görüşme saatleri akşam ve hafta sonlarıdır.)
4. Görüşmeci-yanıtlayıcı etkileşiminin önlenmesi.

Araştırmalarda yanıtlamamaya etki eden bazı faktörler vardır. Bunlar, araştırmanın konusuna, araştırma konusu hakkında halkın kanısına ve bilgisine, veri toplama tekniğine, görüşme süresine, anketör eğitimi ve yeteneğine, anket formunda yer alan soruların formatı ve içeriğine bağlıdır. Bu nedenle, anketin hazırlanmasından raporlanmasına kadar geçen süreçte oluşabilecek hatalar göz önüne alınmalıdır.

Yanıtlamama, araştırma kapsamına alınan birimlerin anket çalışmasına katılmaması sonucu oluşur. Bu nedenle, veri toplama yöntemlerinden anket yöntemine, anketin geçirmiş olduğu aşamalara ve bu aşamalarda hataya neden olabilecek durumlara yer verilmeye çalışılacaktır.

2.1. Anket Yöntemi

Anketin tanımı Türk Dil kurumu tarafından kısa ve öz olarak “Bir konu hakkında soruşturma, araştırma yapmak”, Meydan Larousse’da ise “Aynı soruyu, çeşitli kimselere sorarak bilgi toplamak” şekline verilmiştir.

Birim veya bireylerin önceden hazırlanmış basılı ya da basılı olmayan sorulara yanıt vermesi istenir. Soruların basılı olduğu kağıda soru kağıda ya da anket formu denir.

Anketlerin amacı, tamsayımdan veya üzerinde inceleme yapılan bireylerin oluşturduğu yığından sadece belli bir oranda, çalışmanın amacına bağlı olarak örnek hacmi değişmek üzere çekilen örnekten bilgi elde ederek, kişilerin davranışlarının, demografik özelliklerinin, bilgi düzeyinin ve fikirlerinin ölçülmesidir. Bilimselliği olan anketlerde, örnekler gelişigüzel seçilmez ya da ankete 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 sonuçlar güvenilir bir şekilde büyük yığınlara uyarlanabilmektedir.

Araştırmalar pek çok işlemden geçerler ve bu işlemlerden her biri bir hata kaynağıdır. Bir araştırmanın başından sonuna kadar geçirdiği aşamalarda hata söz konusudur. Analiz aşamasına getirilmiş bir araştırmanın hatalardan ne ölçüde arındığı ve yanıtların güvenilirliği araştırmanın ne kadar iyi düzenlendiğinin bir ölçütüdür (15).

Bu hataların tümünden tamamen kaçınmak mümkün olmasa da, araştırma sonuçlarını etkileyen hataların ölçülmesi ve azaltılması için çaba harcanması gerekir.

Bir araştırmanın geçirdiği aşamalar şöyle ifade edilebilir:

1. Araştırma sorununun teşhisi ve tanımı,
2. Örneklem tasarımı/tamsayım,
3. Veri toplama yöntemi seçimi,
4. Anket formunun hazırlanması,
5. Araştırmanın planlanması ve organizasyonu,
6. Araştırmanın uygulaması,
7. Veri değerlendirme süreci ve analizi,
8. Araştırma raporunun sunumu.

2.1.1. Araştırma sorununun teşhisi ve tanımı

Bir araştırmaya başlamadan önce araştırılacak konunun belirlenmesi gerekir. Konu seçilirken, o konunun bir soruna çözüm getirmesi ya da o alana bir yenilik getirir nitelikte olması, çevre ve toplumun ilgisini çekmesi istenir. Araştırma sorunun öğeleri arasında, araştırma konusu, araştırma hipotezleri ve amaçları yer alır.

Sorunun tanımı aşamasında araştırmanın önemine, çalışmadan elde edilecek sonuçların yararlarına, bilime ve yaşantımıza katkısının neler olduğuna yer verilir. Tanımda şu ilkeler gözetilmelidir (16):

- Problem, alanı konusunda görelî olarak az bilgi sahibi olan birinin anlayabileceği şekilde tanımlanması,
- Spesifik araştırma alanının tanımlanması ve sınırlandırılması,
- Araştırmanın amacının kısaca belirtilmesi,
- “Araştırmanın amacı şudur” gibi açık bir cümlenin bir yerde yer almasına dikkat edilecek şekilde düzenlenmelidir.

Sorun net bir şekilde tanımlandıktan sonraki adım, bir ya da birden fazla hipotezin oluşturulmasıdır. Hipotezler, sorunun yanıtına ilişkin bilgiye dayalı tahminlerdir. İyi bir hipotez, araştırma problemini, bilimsel olarak incelenebilir bir ifadeye dönüştürür.

2.1.2. Örneklem tasarımı/ tamsayım

Araştırma sorunun belirlenmesi, beraberinde yığının belirlenmesini de gerektirecektir. Yığın, ilgilenilen konunun ya da sorunun “taşıyıcıları” olan birimlerdir (14).

Veri toplama işlemi tamsayım veya örnek birimleri üzerinden yapıyor olabilir. Tamsayım araştırma konusu yığın birimlerinin tamamına uygulanan bir çalışma demektir. Nüfus sayımlarında veya yığın genişliğinin az olduğu çalışmalarda tamsayım yapılır. Ancak tüm yığın birimlerinin belirlenemediği, incelenemeyecek kadar çok sayıda birim olduğu, yığın birimlerini inceleme süresi ve maliyetinin çok yüksek olduğu durumlarda tamsayım yapılamaz. Tamsayım yerine yığını temsil edecek nitelikte, yığının yapısına uygun olan bir örneklem yöntemiyle daha küçük genişlikli bir örnek seçilerek yığına ilişkin istenen özellikler belirlenmeye çalışılır (13). Çoğu araştırmada, yığın hakkında bilgi edinme ve yığının özelliklerini öğrenme işlemi örnek üzerinden yapılır.

Örnekleme tasarımı/ tamsayım aşamasında; araştırmanın yapılacağı yer ve bu yerin seçilmesinin nedenleri, yığın hakkında bilgi, örnekleme yapılacaksa örnekleme yöntemi hakkında bilgi verilir.

Burada önemli olan bir nokta, hedef yığın ile araştırma yığını arasında farklılık olabileceğidir. Araştırma konusunun zorunlu kıldığı hedef yığına, olanaksızlıklar, metodolojik imkansızlıklar nedeniyle her zaman ulaşamayabilir. İdeal olan, hedef yığın ile araştırma yığının aynı olmasıdır; bunun sağlanmadığı durumlarda, araştırmacının bunu açıkça belirtmesi gerekir (14).

2.1.3. Veri toplama yöntemi seçimi

Bilimde ve bilgide sağlamlığa ulaşımında en önemli adım, veriyi bilimsel temellere dayalı olarak toplamaktır. Doğru bilgiye iyi toplanmış sağlıklı veriler ile ulaşılır.

Anketler; bilimsel bilgi için en önemli temel kaynaklardır. Anket, bir araştırmada veri toplamak amacıyla posta, telefon, yüz-yüze görüşme yolu ile veya network ağı üzerinden kişilerin görüşlerinin alınması, bu görüşlerin bir araya getirilerek çeşitli açılardan değerlendirilmesidir.

Anketlerle veri toplama, yanıtlayıcıların evlerinde, iş ortamında, sokakta, alışveriş merkezlerinde veya network ağının bulunduğu yerlerde yapılabilir.

En yaygın olarak kullanılan anket yöntemleri şunlardır:

1. Yüz-yüze görüşme anketi
2. Posta anketi
3. Telefon anketi

Anket yöntemi seçiminde, görüşme yapılacak kişilerin yaşam koşulları, araştırma konusu ve kapsamı dikkate alınarak karar verilmelidir.

2.1.3.1. Posta anketi

Posta anketleriyle veri toplama, ya bir konu hakkında mektupla bilgi isteme ya da konuya ilişkin hazırlanan anket formunun gönderilmesi ve takibinde de doldurularak geri yollanmasını sağlamak şeklinde olur (13).

Posta anketlerinde, yanıt oranı düş kırıklığı yaratacak kadar düşük olmaktadır. Posta anketlerinde teşvikler cevaplanma oranını arttırmak amacı ile yaygın olarak kullanılan araçlardan biridir.

Bu amaçla, pek çok araştırmacı, çalışmanın amacını açıklayıp kibarca yardım isteyen, ad soyad açıklanılmayacağı belirtilen bir önyazıyı soru kağıdına ekleyerek yanıt oranını yükseltmeye çalışır. Anket formunun içine pul yapıştırılmış bir zarfın konulması, küçük bir para ödemesin ya da armağan sözü genellikle işe yarar. Yine de yanıtlamayan birimlerinin olması nerede kaçınılmazdır. Bu kişilere ilişkin bilgi edinme, telefon edilmesi veya evlerine gidilmesi gibi daha pahalı yöntemlerle sağlanabilir (17).

Yanıtlayıcılardan anket formlarını kendi başlarına doldurmaları ve posta yoluyla araştırmacıya ulaştırmaları istenir. Özellikle bilimsel araştırmalarda çok yaygın olarak kullanılan bu yöntemin avantaj ve dezavantajları aşağıda sıralanmıştır.

Posta Anketinin Üstün Yönleri

1. Posta anketlerinde kişi sorularla baş başadır. Sorular üzerinde istediği kadar düşünebilir. Cevaplar anketör etkisinden uzaktır, oysa yüz-yüze görüşme anketinde ise anketörün davranışı, soruyu soruş şekli kişinin cevabını doğrudan etkiler.
2. Posta anketleriyle anket formları istenen genişlikte bir örneğe aynı anda uygulanabilir

3. Bazı kişiler kişisel fikirlerini bir yabancıya söylemekten kaçınırlar. Kişiler yazdıklarının gizli kalacağından emin iseler, onlar için yazarak cevap vermek daha uygundur.

4. Gelir düzeyi, eğitim durumu, politik tercih vb. konulara ilişkin bilgiler, yanıtlayıcının anketörden etkilenmesi söz konusu olmadığından posta anketi yöntemiyle daha doğru olarak elde edilir.

5. Yüz-yüze görüşmeye göre daha az masraflıdır. Bu yöntemde anketörlere ihtiyaç yoktur.

6. Posta anketlerinde kişinin evde olmaması ya da anketin uygulandığı yerde bulunmaması nedeniyle anketin yapılamaması söz konusu değildir. Oysa yüz-yüze görüşmede cevaplayıcının evde olmaması anket çalışmasının yapılamaması demektir.

Posta Anketinin Dezavantajları

1. Cevaplama oranı çok düşüktür.
2. Araştırma süresi en uzun olan anket yöntemidir.
3. Araştırmacı soruların doğru kişiler tarafından cevaplanıp cevaplanmadığından ya da soruları cevaplayacak kişinin başkalarının etkisi altında kalıp kalmadığından emin olamaz.
4. Anlaşılamayan soruları açıklama imkanı yoktur.

2.1.3.2. Yüz-yüze görüşme anketi

Yüz-yüze görüşme anketi yönteminde ihtiyaç duyulan veriler, eğitilmiş anketörlerin yanıtlayıcılarla yaptıkları doğrudan görüşmelerle elde edilir. Bu yöntemin diğer yöntemlere göre sağlayacağı avantajlar aşağıda sıralanmıştır.

Yüz-yüze (Kişisel) görüşme anketinin üstün yönleri

1. Okuma-yazma oranının çok düşük olduğu yerlerde uygulanacak en iyi yöntem yüz-yüze görüşme anketi yöntemidir. Ayrıca okuma-yazma bilen kişiler dahi bazı karışık soruları anlayacak düzeyde olmayabilir.
2. Görüşme karşılıklı konuşma şeklinde olduğu için yanıtlamama oranı çok düşüktür.
3. Görüşmede kişi sorulara hemen yanıt vermek zorunda olduğundan dış etkenlerden etkilenmez. Posta anketlerinde bunu sağlamak mümkün değildir.
4. Anketör içten bir hava yaratarak bir çok kişisel, ailevi ve politik sorulara doğru cevaplar alabilir. Posta anketlerinde bu yoktur.
5. Bir cümle, değişik kültür ve sosyal yapıya sahip kişilerce değişik anlamlara gelebilir. Böyle bir yanlış anlama, görüşmecinin kişiye soruyu yeniden daha açık olarak yöneltmesiyle ortadan kaldırılabilir. Posta anketlerinde bunu sağlamak mümkün değildir.
6. Yüz-yüze görüşme, yanıtlamama oranını azaltır.

Bununla birlikte yüz-yüze görüşme yönteminin avantajları yanında bazı dezavantajları da vardır.

Yüz-yüze (Kişisel) görüşme anketinin dezavantajları

1. Maliyeti oldukça yüksektir.
2. Anketörler sağladıkları çok sayıda yarara rağmen, anket sonuçlarının tarafı olmasına yol açabilirler.

3. Cevaplayıcıları evinde bulmak ya da cevaplayıcıların anket için boş zamanlarını yakalamak güç olabilir. Bu gibi aksaklıklar, anket süresinin uzamasına ve maliyetin artmasına yol açar.

2.1.3.3. Telefon anketi

Telefon aracılığıyla veri toplama, bir konu hakkında önceden hazırlanan plan çerçevesinde soru kağıdı ya da basılı olmayan anket formlarıyla kişilerden çeşitli sorulara cevap alma şeklinde olur (13). Telefon anketleri özellikle anket süresinin kısıtlı olduğu durumlarda tercih edilir.

Telefon anketlerinde görüşme öncesi, yanıtlayıcılara ait isim ve adresler mevcutsa, bilgilendirme mektupları gönderilebilir. Böylece beklenmedik telefonların yaratacağı şaşkınlık azaltılarak, anketin tanıtımı, amacı, görüşme süresi aktarılabilir. Bilgilendirme mektupları posta anketi ve yüz-yüze anketlerde de kullanılır. Dolayısıyla bilgilendirme mektupları ile eğitilmiş, deneyimli anketörlerin çalışmada yer alması yanıt oranını arttıracaktır.

Bu yöntemin diğer yöntemlere göre sağlayacağı avantaj ve dezavantajlar aşağıdaki gibi sıralanmıştır.

Telefon anketinin üstün yönleri

1. Kısa zamanda bilgi sağlar.
2. Örnekleme çerçevesi hazır olduğundan örnek seçmek daha kolaydır.
3. Görüşmecinin veya anketörün cevaplayıcı üzerinde oluşturduğu etki, yüz-yüze görüşme yöntemine göre daha azdır.
4. Yüz-yüze görüşme yöntemine göre maliyeti düşük ve daha az kaynak gerektirir.

5. Arařtırmacı yönünden bakacak olursak, çok sayıda farklı kiřilere ulařma imkanı vardır. Ülke içinde veya ülke dıřında kısa sürede çok sayıda bireye ulařabilme imkanı saęlar.

Telefon anketinin dezavantajları

1. Herkesin telefonu bulunamadığından çekilen örneğin yığını temsil etme yeteneęi azalır.
2. Telefon görüşmesiyle elde edilebilecek bilgi sınırlıdır. Ayrıntılı bilgi elde etme olanaęı yoktur.
3. Uzun görüşmeler yapılamaz, kompleks ölçek veya ölçüm türleri kullanılamaz.
4. Telefon anketlerinde Őekil ve resim gibi görsel objeler kullanılamaz.
5. Yüz-yüze görüşme anketine kıyasla yanıt oranı daha düşüktür.
6. Örnek çerçevesi coęrafî olarak ülke sınırları dıřını kapsadığında, yüksek maliyetler söz konusu olacaktır.

Yanıtlayıcılar, soruları tedirgin edici algıladığında yanıtlamama eğiliminde olabilmektedir. Cinsel tercihler veya özel yaşantı gibi konuları içeren çok hassas sorular, özellikle yüz-yüze ve telefon anketlerinde, çoęunlukla yanıtlanmak istenilmemektedir. Bu sorunlardan kaçınabilmek için arařtırmacılar, hassas ve özel olmayan soruların anketörler tarafından sorulması, hassas ve özel sorulara geldiğinde ise yanıtlayıcının kendisinin bizzat tek başına yanıtlanması yönteminin uygulanmasını tercih edebilir. Ancak telefon anketlerinde böyle bir uygulamanın yapılabilmesinin imkanı yoktur (18).

Yüz-yüze ve telefon anketlerinde yanıtlamama ortaya çıkmaktadır. Bunun nedeni:

- Örneğe seçilen birimlere ulaşılamamış olması nedeniyle görüşme sağlanamamıştır,
- Örnek birimlerine ulaşılmış olduğu halde kişiler anketi yanıtlamayı reddetmiştir,
- Örnek birimlerinin fiziksel, zihinsel veya lisan problemlerinin olması nedeniyle bilgi verme yeterliliğine sahip olmamalarıdır (19).

Posta anketlerinde yüksek yanıtlamama durumlarıyla karşılaşmaktadır. Literary Digest anket çalışması, posta anketi yönteminin (1920 ve 30'lu yıllarda) dikkatle kullanılmadığına dair ünlü bir örnektir. Literary Digest başkanlık seçimlerini kestirmek için 10 milyon oy pusulasını postalamıştır. Kestirim hatası 1932 yılında %2 gibi küçük bir değer olmuştur. 1936 yılında ise bu hata %19 gibi büyük bir hata olarak ortaya çıkmıştır. Bu büyük hatanın ortaya çıkması, bir çok kimsenin böyle bir örnekleme yöntemine olan inancının kaybolmasına neden olmuştur. Bu hatanın ana sebebi, yanıtlamayanların yüzdesinin büyük olması olmuştur (20).

2.1.4. Anket formunun hazırlanması

Soruları içeren ve yanıtların üzerine yazılacağı kağıda genel bir ifadeyle anket formu adı verilir (21). Anket formunun düzenlenmesi süreci uygulanan anket yöntemine göre ufak tefek farklılıklar gösterse de temel olarak; kullanılacak soru çeşidi, soru sayısı ve sırası ne olmalı, soruların kapsamı nasıl oluşturulmalı ve soruların ölçülebilirliğinin sağlanması aşamalarını içerir.

Sorular kısa ve basit olmalıdır, alternatifleri açıkça gösterilmelidir. Yanıtlayıcıları anketi doldurmaya güdülemede, ayrıca da başlangıçta onların cevap vermelerini kolaylaştırmada soruların sıralanması önemlidir. Anketin başlangıçta kolay sorularla başlaması, kişiye işi başarabileceği güvenini vereceğinden, soruların sıralanışında, genel sorulardan özel ve ayrıntılı sorulara doğru gitmek önerilen bir yoldur (16).

Anket formundaki soru sayısı araştırmanın amaçlarını en iyi gerçekleştirebilecek sayıda, yani optimum sayıda olmalıdır. Araştırılan konu ile ilgili olmayan, iç bütünlüğe uymayan sorulara yer verilmemesine dikkat edilmelidir.

Anket formu yanıtlayıcıların soruları yanıtlamama veya yanlış yanıtlar verme durumları minimum yapılacak şekilde hazırlanmalıdır. Sorular yanıtlayıcıyı şartlandırmamalıdır. Eğer soru, yanıtlayıcıyı alternatif seçeneklerinden birini seçmesi gerektiği yönünde bir yargıya sevk ediyorsa, yanıtlayıcının şartlanma durumu var demektir.

Anket yönteminde kişilerden belirli bir değişken veya değişkenler bakımından bilgi toplanmaya çalışılır. Kişilerden bilgi toplamada değişkenlerin ölçülebilir olması gerekir. Araştırmalarda ilgilenilen değişkenlerin, fikirlerin sınanması için toplanacak verilerin önemli bir özelliği de ölçülebilir olmasıdır (5). Ölçme çok genel tanımı ile, özelliklerin (değişkenlerin değişim özelliklerinin) gözlenerek sembollerle (rakam, şekil v.b.) ifade edilmesidir (14).

Leigh ve Martin (5)'de, kullanılan ölçek türü ve sorulara verilen yanıt alternatifleri (ölçek noktası) sayısının kişilerin soruları yanıtlamasında etkili olduğu belirtilmiştir.

Ölçek türü, soru zorluğunu ve soru yanıtlamamayı etkileyen bir faktördür (22). Leigh ve Martin (5)'in yürütmüş olduğu çalışmalarda, ölçek türleri için yanıt ve yanıtlamama farklılıklarının ölçek yapısının karmaşık olup olmamasına bağlı olduğu görülmüştür. Çok karmaşık ölçek formatları kullanmanın yanıtlamamaya neden olacağı belirtilmiştir.

Çok sayıda ölçek noktasıyla düşünülen kesinlik, büyük çapta yanıtlamama yanına ve yanıtlayıcı yorgunluğuna yol açabilmektedir. Bununla birlikte az sayıda ölçek noktası kişinin karakteristik özelliğini yansıtmıyor olabilir. Ölçek literatürü incelemeleri sonucunda, Cox (1980) ölçeklerin 5 ile 9 kategoride kullanılmasının uygun olacağını önermiştir (5).

2.1.5. Araştırmanın planlanması ve organizasyonu

Araştırmanın genel amaçları, veri toplama yöntemi, ve kaynakları belirlendikten sonra araştırma faaliyetlerinin planlanması gerekir. Araştırma amacını gerçekleştirmek üzere bir plana dayanmak zorundadır. Planlama ve organizasyon; araştırmanın yönetilmesi, araştırmada kullanılacak materyal ve insan kaynaklarının belirlenmesi ile öntest çalışmasının yürütülmesi aşamasıdır. Bu aşamada, seçilen anket yöntemine göre, anketin yanıtlayıcılara nasıl, ne zaman ve kimler tarafından ulaştırılacağı sorularına cevap verilmeye çalışılır. Yüz-yüze veya telefon anketi yöntemleri uygulanacak ise, alan organizasyonu, anketör seçimi ve eğitimi yapılır.

Posta anketi yönteminde, kişilere bilgilendirme mektubu gönderilebilir. Bu mektubun amacı insanlara anket için seçildiklerini ilerleyen günlerde bir anket formu alacaklarını bildirerek, anketin neden yapılacağı, ankete katılımın önemi hakkında kısaca bilgi vermektir. Böylece posta anketlerinde, çalışmaya katılım oranını arttırmaya çalışılır. Yüz-yüze görüşme veya telefon anketlerinde de kişilere bilgilendirme mektubu gönderilebilmektedir.

Görüşmenin yapıldığı gün, hatta saat, yanıt oranını ve sorulara verilecek cevapların doğasını etkileyen faktörlerdir. Bu nedenle zaman planlamasının yapılması gereklidir. Örneğin evlerde yapılacak görüşmeler için 10:00 ile 21:00 saatleri arası uygundur. Fakat anketlerin hafta içi gündüz saatlerinde yapılması halinde çalışanlar ve öğrencileri evlerinde bulmak güç olacağından, görüşmelerin hafta sonu ya da akşam saatlerinde yapılması daha uygun olacaktır.

Araştırma planlamasında, güvenilir sonuçlara ulaşmada önemli bir adım olan öntest, araştırmanın en son şeklinde yer alacak soruların nasıl daha iyi düzenlenebileceği sorusuna cevap aramak için yapılan bir çalışmadır. Örnek çapından çok daha küçük genişlikte alınan örnek birimlerine öntest uygulanarak, alandan elde edilen geri bildirimlere göre anket formunun yeniden düzenlenmesi sağlanır. Öntestin, yürütülecek çalışmaya olan katkıları şöyledir:

- Örnekleme çerçevesinin yeterli olup olmadığını hakkında bilgi verir,
- Yığın birimlerindeki değişkenliği ve yanıtlamama durumlarını gösterir,
- Olası yanıtlama oranı hakkında bilgi verir,
- Kullanılan veri toplama yönteminin uygunluğuna ilişkin bilgi verir,
- Anket formunun ve anket formundaki soruların yeterliliğini sınıama olanağı verir,
- Talimatların ve görüşmeci eğitiminin yeterliliği konusunda bilgi verir,
- Alan araştırmasının olası maliyeti ve süresine ilişkin veri sağlar,
- Alan-merkez organizasyonunun nasıl sağlanacağı hakkında ipuçları verir.

2.1.6. Alan uygulaması

Alan uygulaması, yığına veya yığından seçilen örnek birimlerine anket formlarının ulaştırılarak yanıt alınmaya çalışılması aşamasıdır. Bu aşama, yüz-yüze ve telefon anketleri için anketörün yanıtlayıcıdan bilgi aldığı, posta anketleri için ise anket formlarının postalanarak, geri dönüşümlerinin takip edildiği süreçtir.

Kişiler, kendilerine gönderilen posta anketlerinin zarfını dahi açmayabilir, açıp bir kenara atabilir, doldurmaya başladıktan sonra sıkılarak yarıda bırakabilir ya da anketi doldurmalarına rağmen geri göndermeyebilirler. Yanıtlayıcıların postalanan anket formlarını ellerine almasından, doldurdukları anketin geri yollanmasına kadar geçen zaman içerisinde ankete olan ilgilerinin sürmesi gerekmektedir. Bu nedenle, anketler gönderildikten bir hafta 10 gün sonra, örneğe alınmış tüm birimlerden anketi cevaplayıp gönderenlere teşekkür edildiği, henüz göndermemiş olanlara hatırlatmanın yapıldığı bir kart gönderilmeli, böylece anket çalışmasına olan katılım takip edilmelidir.

Yüz-yüze görüşme, posta ve telefon anketleri uygulamasında araştırmacının tamamen inisiyatifi dışında kalan bir yanıtlayıcı bileşeni vardır. Yanıtlayıcı soru kağıdında yer alan sorulara yanıt veren kişidir. Dedek (23), yanıtlayıcının cevap verirken etkilenmiş olduğu faktörleri şöyle vermiştir.

1. Örnekleme tasarımından etkilenir. Tasarımda uygun olmayan yanıtlayıcı birim tanımlanmamış ya da kapsam dışındaki bir birim kapsama dahil edilmiş (ulaşılabilir kişi doğru kişi değildir) veya dahil edilmemiş olabilir.
2. Soru kağıdının tasarımından etkilenebilir. Soru kağıdının soru akışı doğru tanımlanmamış olabilir ya da sorularda yer alan seçenekler doğru tanımlanmamıştır.
3. Yanıtlayıcı kendi sosyal, demografik ve ekonomik yapısından etkilenir.
4. Sosyal çevresinden etkilenir.
5. Yanıt verdiği kişinin yapısından etkilenir.

Birinci ve ikinci maddeler araştırma tasarımını yapan kişilerce öntest çalışmalarından tespit edilerek denetim altına alınabilir. Üçüncü ve dördüncü maddeleri fiziki olarak değiştirmek mümkün değildir. Kişinin yaş, cinsiyet ya da geliri değiştirilemez zaten büyük bir ihtimalle ölçülmek istenen değişkenler de bunlardır. Kişinin sosyal çevresi kişinin davranışlarında önemli paya sahiptir. Kişi araştırmayla ilgili olarak çevresi nedeniyle yanıt vermeme eğiliminde olabilir ya da çevresince yanlış değerlendirileceğini düşünerek gerçeği söylememe eğiliminde olabilir. Beşinci madde ise araştırma tasarımını yapan kişilerce anketörün eğitimi ile denetim altına alınabilir.

Francis ve Busch (24), üç başkan seçiminin yapıldığı bir çalışmada soru yanıtlamama miktarının kestirimlerini incelemiştir. Politik faktörlerde araştırma konusu kapsamının en güçlü kestirici olduğu görülmüştür. Politikayla ilgilenen kişilerde çok düşük soru yanıtlamama seviyesi olduğu görülürken, politikayla ilgilenmeyen kişilerde politik görüş sorularında yanıtlamama durumları olduğu gözlenmiştir.

Faulkenbery ve Mason (25), "Hiçbir Fikrim Yok", "Bilmiyorum", "Fikrim Var" yanıt ilişkilerini, elektrik jeneratörlerinde yel değirmeni kullanılması hakkında kişilerin görüşlerinin ne olduğunun sorulduğu bir soruda araştırmıştır. "Hiçbir Fikrim

Yok”, “Bilmiyorum” yanıtlarını veren kişilerin, “Fikrim Var” yanıtını veren kişilere kıyasla konu hakkında oldukça az bilgiye sahip olduğunu tespit etmişlerdir.

Yanıtlamama faktörünün minimize edilebilmesi için, araştırmacılar konu hakkında en azından yaklaşık bilgisi olan kişileri seçebilir ve basit sorular sorabilir (5).

Pek çok durumda tüm çabaların boşa gitmesine neden olabilecek yanıtlamama sorununun yegane çözümü; insanların anket formunu doldurarak elde edecekleri kazancın harçayacakları çabaya değeceğine ikna etmektir. Ancak burada kazanç yalnızca maddi kazanç olarak düşünülmemeli, kişilerin topluma sağlayacağı katkı, profesyonel bir çalışmaya katılma, kendini önemli hissetme, fikirlerini açıklayabilme, sesini duyurabilme fırsatı bulma v.b. gibi duygular da yanıtlayıcıya hissettirilmeli ve çalışmaya katılım sağlanmalıdır.

2.1.7. Verilerin işlenmesi ve analizi

Veriler toplandıktan sonra, verilerin gözden geçirilmesi ve işlenmesi sonrasında da analize hazır hale getirilmesi gerekir. Bunun için şu işlemler yapılır: Edit, Kod, Veri girişi.

Edit işlemi ile doldurulan her bir anket formu ayrı ayrı incelenerek her bir görüşmenin doğruluğu ve tam olma durumu gözden geçirildikten sonra, veriler kodlamaya hazır hale getirilir. Bu işlem veriler toplandıktan hemen sonra, uygulama aşamasının sonunda yapılır.

Kod işlemi, anket formundaki cevapları harf veya sayı şeklindeki sembollerle tanımlamayı ve sınıflandırmayı gerektiren bir süreçtir. Kod aşamasında oluşabilecek hatalar, kodlamayı yapan kişiden kaynaklanan hatalardır. Veri girişi, kodlama işleminin tamamlanması sonunda verilerin bilgisayar ortamına aktarılarak analizi yapılacak hale getirilmesi aşamasıdır.

Bu süreçlerden geçen veri çeşitli istatistiksel yöntemlerle analiz edilir. Veri analizinin amacı, eldeki problemi çözmeye yardımcı olacak bilgiyi sağlamaktır.

2.1.8. Araştırma raporunun sunumu

Problem ve amaç ortaya konulduktan sonra, analizler sonucunda elde edilen bilgiler rapor haline getirilerek sunulur.

Buraya kadar bir araştırmanın geçirdiği aşamalar üzerinde durulmuştur. Anket çalışmasında amaç hedef yığında bulunan tüm birimlerden veya yığından tesadüfi olarak seçilen örnek birimlerinden yanıt almaktır, yanıtlamama durumu istenilmeyen bir durumdur.

Örnekte bulunan birim veya bireylerin bazılarının ankete yanıt vermemesi ya da kısmen yanıt vermesi yanıtlamama sorunun çıkmasına neden olur. Sonuçların yanlışlık göstermemesi için, bireylerin araştırmaya hiçbir etki altında kalmadan katkıları beklenir (4).

2.2. Araştırmalarda Yanıtlamama

Araştırmalarda, birim veya bireylerin çeşitli nedenlerle çalışmaya katkıda bulunmama ya da kısmen katkıda bulunma durumlarıyla sıklıkla karşılaşmaktadır. Araştırmalarda yanıt oranı veri kalitesinin kritik bir belirleyicisidir (26). Düşük yanıt oranı bulguların geçerliliğine ve genelleştirilmesine zarar vereceğinden, bir sorun olarak kabul edilen yanıtlamama sorununun giderilmesi için çalışmalar yapılmalıdır (27).

Araştırmalarda yanıtlamamayı; sırasıyla yanıtlamamanın belirlenmesi, yanıtlamama türleri, yanıtlamama kaynakları ve yanıtlamama etkileri başlıkları altında incelenecektir. Burada, yanıtlamamanın ne olduğu, nedenleri, nelere yol açtığı ele alınmaya çalışılmıştır. Ayrıca, yanıtlamama sorunun hatalar içindeki yerinin ne olduğunu daha iyi anlayabilmek amacıyla, alt başlık altında yanlar ve örnekleme dışı hatalara da tezde yer verilmiştir.

2.2.1. Yanıtlamamanın belirlenmesi

Särndal ve Wretman (28), anket çalışmasının amacını q sayıda y_1, y_2, \dots, y_p çalışma değişkenleriyle örnek birimlerinin gözlenmesi olarak tanımlamıştır. Anket araştırmalarında tam yanıt olması, verilerin toplanıp derlenmesi işleminden sonra, p vektöründe tüm $i \in s$ için gözlenen değerlerin mevcut olması, $n \times p$ boyutlu veri matrisinde hiç bir değer kayıp değildir demektir. Tüm diğer durumlarda ise, yanıtlamama vardır demektir. Burada $i=1, \dots, n$ ve $j=1, \dots, p$ iken, y_{ij} : i . birim için y_j değişken değeri, p : değişken sayısı ve n : örnek genişliğidir.

$n \times p$ veri matrisinde tüm birimler için p değişken değerlerinin mevcut olması tam yanıt olması demektir. $n \times p$ veri matrisinde kayıp değer yoktur, diğer tüm durumlarda ise yanıtlamama vardır demektir.

Aşağıda tam yanıt olması ve yanıtlamama durumları için çeşitli örnekler farklı veri kümeleriyle gösterilmiştir.

$$\begin{array}{c}
 \text{Değişken } 1, \dots, p \\
 \left(\begin{array}{cccc}
 y_{11} & y_{12} & \dots & y_{1j} \dots y_{1p} \\
 y_{21} & y_{22} & \dots & y_{2j} \dots y_{2p} \\
 \cdot & \cdot & \cdot & \cdot \\
 y_{i1} & y_{i2} & \dots & y_{ij} \dots y_{ip} \\
 \cdot & \cdot & \cdot & \cdot \\
 y_{n1} & y_{n2} & \dots & y_{nj} \dots y_{np}
 \end{array} \right)
 \end{array}$$

Şekil 2. 1. Tam yanıt olması durumu

		Değişken 1, , p				
Birimler 1, , n	y ₁₁	?	y _{1j}	y _{1p}
	y ₂₁	?	y _{2j}	y _{2p}

	y _{i1}	?	y _{ij}	y _{ip}

.	
y _{n1}	?	y _{nj}	y _{np}	

Tüm birimler 2.değişkeni yanıtlamamıştır.

Şekil 2. 2. Yanıtlamama durumu

		Değişken 1, , p				
Birimler 1, , n	?	y ₁₂	y _{1j}	?
	y ₂₁	?	y _{2j}	y _{2p}

	y _{i1}	y _{i2}	y _{ij}	?

.	
y _{n1}	?	y _{nj}	y _{np}	

Birimler bazı soruları yanıtlamamış veya kayıp değer vardır.

Şekil 2. 3. Yanıtlamama durumu

		Değişken 1, , p				
Birimler 1, , n	y ₁₁	y ₁₂	y _{1j}	y _{1p}
	?	?	?	?

	y _{i1}	y _{i2}	y _{ij}	y _{ip}

.	
y _{n1}	y _{n2}	y _{nj}	y _{np}	

İkinci birim soruların tümünü yanıtlamamıştır.

Şekil 2. 4. Yanıtlamama durumu (29)

Veri toplama ve derleme işleminden sonra, veri kümesinde bir ya da birden fazla y_{ij} , $i=1, \dots, n$ ve $j=1, \dots, p$ değeri mevcut değil ise $n \times p$ veri matrisinde değerlerin olması gereken yerlerde “boşluklar” olduğu gözlenir, $n \times p$ matrisinde y_{ij} gözelerinde boşluklar olacaktır. Bu durumda yanıtlamama durumuyla karşı karşıyayız demektir.

Bir anket çalışmasında tam yanıt olması nadir bir durumdur, y_{ij} kayıp verilerinin olması için çeşitli nedenler vardır. Bir posta anketi çalışmasında, kişilere postalanan anket formu geri gönderilmemiş veya gönderilmiş ancak tam doldurulmamış olabilir. Yüz-yüze görüşme anketi çalışmasında, birimler bazı soruları veya tüm soruları yanıtlamayı reddetmiş, kişilere ulaşılamamış, yanıtlar anket formuna kaydedilmemiş veya veri girişi yapılmamış olabilir. Böyle daha başka nedenlerden ötürü yanıtlamama durumu ortaya çıkabilmektedir.

2.2.2. Yanıtlamama türleri

Yanıtlamama, örneğe alınan bazı birim veya bireylerin ankete yanıt vermemesi ya da kısmen yanıt vermesi durumudur. Araştırmalar, yanıtlamama türlerinin etkilerini ve yanıtlamama durumlarını minimum yapacak, yanıt vermeyen birimler ile yanıtlayan birimler arasındaki farklılıklardan ortaya çıkan hataları en aza indirgeyecek şekilde tasarlanmalıdır. Yanıtlamama, birim yanıtlamama ve soru yanıtlamama olarak iki ayrı grupta incelenebilir.

Birim yanıtlamama; örnekleme çerçevesinde bulunan bir veya birden fazla kişi veya birim anket çalışmasına katılmadığında ortaya çıkar. Yukarıda verilmiş olan şekillerden Şekil 2.4 birim yanıtlamama durumunu göstermektedir. Soru yanıtlamama; örneğe seçilen bir veya birden fazla kişi anket çalışmasına katılıyor ancak en az bir değişkeni veya soruyu yanıtlamıyorsa ortaya çıkar (30). Şekil 2.2 ve 2.3 soru yanıtlamama durumunu göstermektedir.

Çalışmada yalnızca birim yanıtlamama durumu incelenmiştir. Birim yanıtlamama, örneğe alınan birimlerin bazılarıyla gözlem yapılamaması ve bu birimlere ilişkin hiç veri toplanamamış olması durumudur. Birim yanıtlamamayı “boş” veri vektörüyle

göstermek mümkündür (y_i : i. yanıtlamayan birimine ilişkin gözlem değerlerini gösteriyorken, $y_i = (_ , _ , \dots, _)$ olur).

Birim yanıtlamama, U.S. Census Bureau Survey of Income and Program Participation (SIPP)'da ev halkı yanıtlamama ve kişi (veya birim) yanıtlamama olarak iki ayrı grupta incelemiştir (31).

Ev halkı yanıtlamama: Anketi yanıtlama yeterliliğine sahip ev halkı veya birimlerin (kurum, şirket gibi) tümü ankete katılmayı reddediyor ya da bütün girişimlere rağmen tümüne ulaşamıyorsa ortaya çıkar. Anketi yanıtlama yeterliliğine sahip ev halkı bireylerinden herhangi birine ulaşılması yanıt alınması için yeterli olmaktadır.

Kişi yanıtlamama: Anket çalışmasının konusu itibariyle yalnızca örneğe seçilen kişiyle görüşme yapılarak veri toplanması gerektiğinde ortaya çıkan bir yanıtlamama durumudur. Bu, kişilerin işleri veya meslek alanlarındaki çalışmalarıyla ilgili bilginin alınacağı bir anket çalışması olabilir. Kişi yanıtlamama, kişiye ulaşamaması veya kişinin anketi yanıtlamayı reddetmesi durumunda ortaya çıkar. Kişi yanıtlamamanın, ev halkı yanıtlamama durumuna kıyasla daha yüksek yanıtlamama oranına sahip olması olasıdır.

Çalışmamızda kişi ve ev halkı yanıtlamama birim yanıtlamama olarak bir başlık altında ele alınmıştır.

2.2.3. Yanıtlamama kaynakları

Yanıtlamama kaynakları, örneğe alınan bazı birimler için gözlemleri elde etmede başarısızlığın kaynaklarını belirtir. Gelecekte yapılacak çalışmalarda sakınabilmek ve araştırmalara olabilecek etkilerinin tahminlerini yapabilmek için, yanıtlamama kaynaklarının anlaşılabilir olarak kontrolünün yapılması ve azaltılması faydalı olur (6).

Yanıtlamama kaynaklarının incelenmesi ve değerlendirilmesi araştırma sonuçlarının yorumuna ışık tutması açısından önemlidir. Yanıtlamama kaynaklarının iyi

sınıflandırılması anket araştırmanın türüne ve durumuna bağlıdır. Aşağıda verilecek sınıflandırma yüz-yüze görüşme anketleri temel alınarak yapılmış olup, diğer anket türlerine de bir öneri oluşturacak şekilde düzenlenmiştir.

2.2.3.1. Evde kimse yok

“Evde kimse yok” kategorisi içerisinde kişi gerçekte evde ikamet ettiği halde, araştırma ekibinin alanda çalıştığı süre içerisinde evde olmayanlar alınmıştır. Kişi geçici olarak evde yok, gezmeye gitti veya çalışıyorsa evde olmama durumudur.

Yanıtlayıcının görüşme yapılmasını reddetmesi değil, görüşmeyi erteleme durumu belirtir. O anda görüşmeci çok yoğun, yorgun veya hasta olabilir; ya da hava koşulları kişilere ulaşmaya engel olabilir. Böyle durumlarda görüşme yapılmak üzere kişi yeniden aranmalı veya kişiye yeniden ulaşılmaya çalışılmalıdır (32).

Bu, yoğun telefon görüşmelerinin yapılacağı, posta anketlerinde ise gecikmelerin ve isteksizliklerin yaşanacağı anlamına gelir.

2.2.3.2. Yöre dışında

Çalışmanın yapıldığı tarihlerde yanıtlayıcıların evden geçici olarak uzak olması durumudur. Anket çalışmalarında mevsimsel değişimler dikkate alınmalıdır. Özellikle yaz aylarında, kişilerin izinli olduğu zamanlarda görüşmelerde aksaklıklar olabilmekte, kişi tatile çıkmış veya iş seyahatinde olabilmektedir (7).

2.2.3.3.Redler

Çeşitli nedenlerle araştırmaya katkıda bulunmayı istememe durumudur. Yanıtlayıcıların özellikleri, kültür farklılıkları, sosyal sınıfları ve demografik yapıları farklı olabilmektedir. Araştırmanın kapsamı, anket formundaki soruların içeriği ve sırası, anketörler anket çalışmasının reddedilmesini etkileyen önemli faktörlerdir.

Değişik kültürlerde gelir, inanç, ırk ve politik görüşlerin tartışılması çok güç veya kolay olabilir.

Bu türden bir çok faktör bulunabilmektedir. Kişilere hasta iken veya uygun olmayan bir saatte ulaşıldığında görüşmenin reddedilmesi durumu söz konusu olmaktadır. Bu durumda, sonrası için görüşme yapılamak istenildiği söylenilmeli, kişilere tekrar ulaşılmaya çalışılmalıdır.

Tüm girişimlere rağmen redler oluyorsa, böyle kişileri genel bir kategori de “Elde edilemez” olarak sınıflandırmak mümkündür.

2.2.3.4. Yetersiz ya da ehliyetsiz

Anket çalışması süresince kişinin hasta olması, fiziksel veya zihinsel sorunlarının olması yanıt vermesine engeldir. Bazı anketlerde kişilerin lisan sorunun olması, okuma-yazma bilmemeleri yanıt alamama durumunu doğurabilir.

2.2.3.5. Ulaşılamayanlar (bulunamadı)

Bu kategoride ki kişiler görüşme girişiminde bulunulduğu halde ulaşılamayan kişilerdir. Anket süresince yanıtlayıcıların yaşadıkları bölgelerin tehlikeli, bölgelerde doğal afetler, sel, deprem veya çok ağır hava koşullarının olması çalışmayı olumsuz etkileyen faktörlerdir. İşleri nedeniyle veya başka sebeplerden ötürü sürekli seyahat etmek durumunda olan kişilere ulaşılması çok güçtür. Böyle yanıtlayıcılar belirlenemez ve takip edilemez.

Ulaşılamayan veya bulunamayan kişiler, bu kişilerden bilgi alınamadığı için yanıtlanmayanlar kategorisinde sınıflandırılır (6).

2.2.3.6. Kayıp listeler

Anket formlarının işyerinde zarar görmesi, posta anketlerinin gönderildiği halde kaybolması, bazı anket formlarının düşük kalitede veya dürüst olmayan yollarla doldurulmuş olması, anketlerin kullanılmamasına neden olur.

Görüşme öncesi kaybedilen veya unutilan listeler de kayıp listeler olarak bu kategoride yer alırken yanıtlamamanın bir kaynağını oluşturmuş olur.

2.2.4. Yanlar ve örnekleme dışı hatalar

Araştırmalar pek çok işlemi içerirken, bu işlemlerin tümü hataya neden olur. Yanıtlamama hatalarına geçmeden önce yanlar ve örnekleme dışı hatalar incelenerek, yanıtlamamanın hatalar içindeki yeri verilmeye çalışılacaktır.

Araştırma konusu değişkenin, gerçek değeri ile araştırma istatistiği arasındaki fark toplam hata olarak adlandırılır. Araştırma istatistiği gerçek değerden iki nedenden farklı olabilir. Bilgi tam sayımdan değil, örnekleme dayalı araştırmadan elde edildiği için: Bu tür hatalar örnekleme hatalarıdır. Ölçüm sürecinde hatalar yapıldığı için: bu tür hatalar örnekleme dışı hatalardır (6).

Araştırmalarda toplam hata; örnekleme hataları ve örnekleme dışı hatalar olarak ikiye ayrılabilir. Yığının yalnızca bir altkümesine ilişkin bilgi olması yüzünden ortaya çıkan hatalara örnekleme hataları denilir. Parametre değeri ile ölçümlerden elde edilen tahminler arasındaki fark da örnekleme hatası olarak da tanımlanır (17). Bu hataların etkileri azaltılabilmekle birlikte hiçbir zaman tam olarak ortadan kaldırılamaz. Diğer bir ifadeyle, hedef yığında bulunan tüm birimler ankete dahil edilmediği sürece örnekleme hatası kaçınılmazdır.

Uygulamalı çözümlenelerde, kullanılan örnekleme süreciyle ilişkisi olmayan hatalara örnekleme dışı hatalar adı verilir (17). Üç tür örnekleme dışı hata vardır:

kapsam hatası, ölçüm hatası ve yanıtlamama hatası. Anketin uygulama aşamasında, kayıt tutma, kodlama, bilgi işlem ile yorum aşamalarında yapılan hatalar örnekleme dışı hatalardır. Bu tür hatalar, hem örnek araştırmalarında hem de tam sayımlarda vardır ve tam sayımlarda daha büyük çaplı olmaktadır (15).

Örnekleme dışı hataları tanımak ve çözmek için genel bir süreç yoktur. Önerilebilecek başlıca yol; araştırmacının, hedef yığını belirleme, anket formunu düzenleme, yanıtlamayanların etkisini en aza indirmeye çalışma gibi konularda özen göstermesidir (17).

Örnek çapı büyüdükçe, örnekleme hataları azalırken, örnekten elde edilen tahmin değeri parametre değerine (gerçek değere) 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 (32).

İyi planlanmış ve yürütülmüş bir örnek araştırmasının toplam hatası tam sayıma kıyasla çok daha küçük olabilir. İyi bir anket çalışması için öncelikle iyi bir çerçeve (yığının listesi) hazırlanarak kapsam hatasından kaçınılmalıdır. İkinci olarak, örnekleme hatasını büyük ya da küçük örnek seçerek kontrol etmelidir, burada verilmesi gereken karar ne kadarlık hataya müsaade edilebileceğidir. Üçüncüsü ise, ölçüm hatasına yol açabilecek önyargılı ve karmaşık sorulardan kaçınılmalıdır. Ölçüm hatasının boyutu tespit edilememekle birlikte, bu hata, soruların dikkatli hazırlanması ve önceden test edilmesiyle azaltılabilir.

Son olarak, soruları mümkün olan en yüksek yanıt oranına ulaşılacak şekilde tasarlamalıdır. Bu şekilde yanıtlamama hatasının gerçekleşmesi durumu azalacaktır. Ankete dahil olan kişiler ve kullanılan anket yöntemine bağlı olarak, %50-60'ın altındaki yanıtlama yüzdesinin sebebi mutlaka araştırılmalıdır ve soruları yanıtlayanların, yanıt vermeyenlerden farklı olup olmadıkları incelenmelidir.

Bu hataların tümünden kaçınmak mümkün olmasa da, anketin hazırlanması, tasarımı, uygulanması ve değerlendirilmesi aşamalarında oluşabilecek bu hatalar göz önüne alınarak çalışma yürütülmelidir.

Araştırmalarda yapılan sistematik hatalara *yan (bias)*, tesadüfi kaynaklı hatalara ise *değişken hataları (random error)* adı verilir. Örnekleme ve örnekleme dışı hatalar tesadüfi veya sistematik kaynaklı olabilir. Toplam hatayı tesadüfi ve sistematik kaynaklı hatalar şeklinde ikiye ayırarak incelemek de mümkündür (6).

Buna göre hatalar dört başlık altında sınıflandırılabilir (6).

- | | | |
|------------------------------------|---|-----------------------------|
| 1. Değişken örnekleme hataları | } | Tesadüfi kaynaklı hatalar |
| 2. Değişken örnekleme dışı hatalar | | |
| 3. Örnekleme yanlılığı | } | Sistematik kaynaklı hatalar |
| 4. Örnekleme dışı yanlılık | | |

Örnekleme teorisinde, yaygın olarak kabul edilmiş olan bir model, değişken hatası ve yanın birleşimini toplam hata içinde toplamıştır. Bu model, hata kareler ortalamasının (HKO) kareköküne eşittir. Bir tahmin ediciyle gerçek değeri arasındaki farkın karesi, beklenen hata kareler ortalaması olarak adlandırılır. μ yığın ortalaması (gerçek değer) için hata kareler ortalaması aşağıdaki gibidir.

$$HKO = E [\bar{x} - \mu]^2$$

$$HKO = E [\bar{x} - E(\bar{x})]^2 + E [E(\bar{x}) - \mu]^2$$

$$HKO = \text{Varyans}(\bar{x}) + (\text{Yanlılık})^2$$

$$TOPLAM HATA = \sqrt{DH^2 + Yan^2} ; \quad DH: Değişken hatası$$

Değişken hataları (DH), yalnızca örnekleme hatalarından dolayı ortaya çıkmışsa, DH^2 örnekleme varyansına eşit olur. Örnekleme dışı hata olmaması durumunda, örnekleme yanlılığı yoksa, örnek tahmini parametre tahminine eşit demektir (32).

Toplam hatanın diğer bileşeni yan, araştırmalardan elde edilen ortalama değerlerin gerçek yığın değerinden sapmasıdır. Yanların, bir araştırmanın toplam hataları üzerinde önemli etkileri var ise, toplam hata yalnızca değişken hatalarıyla düşük tahmin edilmiş olacaktır. Yanlar “gerçek değer” ile beklenen değer arasındaki farklılığı gösterir. Buradan toplam yan;

$$(\sum B_g = E(\bar{x}) - \mu)$$

şeklinde gösterilebilir. Buna karşın, değişken hataları; tahmin ile tahminin beklenen değeri arasındaki farklılığın $(\bar{x} - \mu)$ kaynağını ölçer.

Yetersiz bir örnekleme çerçevesinin kullanılması, örnek seçim sürecinin yanlış uygulanması, verilerin toplanma ve değerlendirilme işlemlerinin hatalı olması vb. gibi nedenler yanlı tahminlere yol açabilir (10).

Çoğu yan örnek genişliğinin artırılmasıyla azaltılamaz, yalnızca bazı işlemlerin kalitelerinin artırılmasıyla azaltılabilir. Değişken hatalarının azaltılması ise bunun tam tersi olarak örnek birimleri, gözlemler veya gözlemcilerin sayısının artırılmasıyla sağlanır.

Araştırma sonuçlarını etkileyen hataların ölçülmesi ve azaltılması için çaba harcanması gerekir. Groves (1), hataların azaltılmasının;

- İstenen bilginin yüksek kalitede alınmasını sağlayacak soruların hazırlanmasıyla,
- Yığını en iyi temsil edecek örnekleme yöntemi seçilmesiyle,

- Yanıtlayıcılar üzerinde anketör etkisinin azaltılmasıyla,
- Örneğe alınan birimlerin tamamından yanıt alınması veya yanıtlanmamanın minimum yapılmasıyla gerçekleştirilebileceğini belirtmiştir.

2.2.5. Yanıtlanmamanın etkileri

Örnek çapı ve yanıt oranı, örnek araştırmasının kalitesinin belirleyicileridir (33). Yanıt oranı %100 oluğunda ancak parametrelerin yansız tahminleri yapılabilir. İstatistikçiler ve araştırmacılar, yanıtlanmamanın tahminlerin kalitesini düşürdüğünü belirtmektedirler. Yanıtlanmama araştırma bulgularını etkilediğinde önemli bir hata bileşeni olarak karşımıza çıkmaktadır (34).

Groves ve Lyberg (19), araştırmalarda yanıtlanmamanın üç nedenden dolayı sorun olduğunu belirtmişlerdir. Birincisi, yanıtlanmayanların yanıtlayanlardan farklı olması nedeniyle, yalnızca yanıtlayanlara dayalı yapılan parametre tahminlerinin yanlı olmasıdır. İkincisi, yanıtlanmayan birimlerinin varlığı, örnek çapında azalmalara neden olur. Son olarak ise yanıtlanmama araştırma maliyetini artırır.

Yanıtlanmama oranı önemli boyutta ise, araştırmada ek örnekleme hatası ve örnekleme dışı hatalara yol açar. Örnekleme hatası doğar çünkü örnek büyüklüğü azalmıştır. Örnekleme dışı hatalar oluşabilir; çünkü örnek alınan yığın aslında ilgilenilen yığın değildir. Elde edilen bulgular, yanıt vermeye istekli yığından tesadüfi seçilmiş bir örnek gibi görülebilir. Eğer yanıtlanmayanlar yanıtlayanlardan farklı ise yapılan tahminler yanlı olacaktır (17).

Yanıtlanmamanın araştırma sonuçları üzerinde iki önemli etkisi vardır, örnek büyüklüğünü azaltır ve yanlılık doğurur.

2.2.5.1. Yanıtlamama oranı

Örnekleme kalitesinin ölçümünde sorulacak ilk soru; örnek çapının ne olduğudur, bir başka daha az sorulan ancak bu soru kadar önemli olan diğer bir soru ise; yanıt oranının ne olduğudur. Örnekleme hatasının ölçümünde, örnek çapı elbette ki önemlidir. Ancak örnek çapından bağımsız bir örnekleme dışı hata olan yanıtlamama yanının belirlenmesinde, seçilen örnek birimlerinden ne kadarından yanıt alındığı çok önemlidir (33).

Örnek büyüklüğündeki azalma araştırma sonuçlarının güvenilirliğini azaltacağı gibi, örnekleme varyansını da artırır. Dolayısıyla yanıtlamama oranının düşük olması istenir (6).

Araştırmalarda yanıtlamamaya etki eden bazı faktörler vardır, bu faktörler yanıtlamama oranını doğrudan veya dolaylı olarak etkileyebilmektedir. Araştırmanın konusu, bu konuda halkın kanısı, veri toplama tekniği, görüşme süresi, anketörün yetenek ve tavırları, yörede seçilen diğer hane halklarının yanıt oranı ve hatta mevsimler yanıtlamamaya etki eden faktörlerdir.

Yanıtlayanların ve yanıtlamayanların oranı yanıtlamama hatasının açıklanmasında kullanılabilir (1).

Yanıtlamama oranı araştırmadan araştırmaya ve araştırma organizasyonundan araştırma organizasyonuna çok büyük farklılıklar gösterir. Bu oranlar zaman içinde, hatta aynı araştırmanın tekrarlarında dahi değişim gösterir. Örnek için yanıt oranı aşağıdaki gibi tanımlanabilir.

$$p = \frac{r}{n}$$

Burada r yanıt verenlerin sayısı, n ise örnek çapıdır. Örnek için yanıtlamama oranı,

$$q = 1 - p$$

olarak tanımlanır. Bu ölçüm bir araştırmanın başarısının göstergesidir. q ne kadar düşük ise yürütülen araştırma o kadar iyi demektir (35). The Panel On Incomplete Data In Sample Surveys (1983)'de, yanıtlamama ve red oranlarının artış gösterdiği belirtilmiştir. İki U.S. fon araştırmasında 1968'den 1980 yılına kadar yanıtlamama oranları karşılaştırılmış ve red oranlarının zaman içerisinde artış gösterdiği görülmüştür.

Örnekleme dahil edilen kişilerin önemli bir kısmı ile görüşülememesi, örneğin, %20'lik bir yanıt oranı, yanıtlamama hatasına yol açar. Hedef yığının yalnızca %20'sinden alınan sonuçların, tüm yığının görüşlerini yansıtmaması, bazı özel durumlar dışında çok güçtür. Bununla birlikte yüksek yanıt oranının, yanıtlamama hatasını tamamen önleyeceği de düşünülmemelidir. Örneğin, ülkemizde yapılan 2000 yılı nüfus sayımında, toplumun ortalama %94'ü ile görüşülmüş olmasına rağmen, kalan %6'lık bölümün çok farklı özellikler göstermesi (varoşlarda yaşayan kişiler) yanıtlamama hatasına yol açmıştır (36).

2.2.5.2. Yanıtlamama yanlılığı

Yanlılık gerçek yığın değeri ile beklenen değer arasındaki farktır. Yanlılığın diğer bir tanımı, bir sonucun istatistiki olarak temsil gücünü sistematik kaynaklı sebepler nedeniyle kaybetmesi olarak verilebilir (7).

Yanıtlamama yanlılığı, araştırma konusu değişken değer(ler)inin, yanıtlamayanlar ve yanıtlayanlar için farklılığı nedeniyle kendini gösterir. Yanıtlamayanlar, yanıtlayanlardan çok farklı ise, yanlılığın etkisi daha da ciddi kendini gösterir. Bu konuda yapılan araştırmalar, bu iki grubun birbirinden farklı olduğunu, bu grupların hiç bir koşul altında aynı olduğu varsayımının yapılamayacağını göstermektedir (7).

Örneğin eğitim düzeyi ile politik görüş arasındaki ilişkinin araştırıldığı, teknik terimler içeren bir posta anketinin toplumun geneline uygulandığını varsayalım. Bu tür bir anketi eğitimsiz kişiler anlamakta güçlük çekeceklerinden, toplumun bu

kesimi için yanıt oranı düşük olacaktır. Dolayısıyla, geri dönen anketlerin büyük bir kısmı eğitilmiş kişiler tarafından doldurulacağından, elde edilen sonuçların yanlış olması kaçınılmazdır.

Yanıtlamama yanlılığının araştırılmasının önemini verebilmek için, Cochran (10), Kish (6), Moser ve Kalton (7)'un bu konudaki klasik örneğini ele alınmıştır: Yanıtlayan ve yanıtlamayanları iki ayrı tabaka olarak düşünülün. Yanıtlayanlar tabakasına ait birimlerin sayısı N_1 , yanıtlamayanlar tabakasına ait birimlerin sayısı N_2 ile gösterildiğinde, yanıt oranı;

$$W_1 = N_1 / N,$$

ve yanıtlamama oranını;

$$W_2 = N_2 / N$$

şeklinde tanımlanabilir ($W_1 + W_2 = 1$). Burada, yanıtlamayanların varlığında yığın ortalaması aşağıdaki gibidir.

$$\mu = \frac{N_1}{N} \mu_1 + \frac{N_2}{N} \mu_2$$

- N : Yığın çapı
- N_1 : Potansiyel yanıtlayıcı olan yığın birimlerinin sayısı
- N_2 : Potansiyel yanıtlamayan yığın birimlerinin sayısı ($N_2 = N - N_1$)
- μ : Yığın ortalaması
- μ_1 : Yanıtlayanların yığın ortalaması
- μ_2 : Yanıtlamayanların yığın ortalaması
- W_1 : Yanıt oranı
- W_2 : Yanıtlamama oranı

Yığından basit tesadüfi örnekleme yöntemiyle seçilen bir örnekte, yanıtlamayan birimlerin olduğu ve yanıtlamayan birimlerinden veri toplamak için herhangi bir girişimin yapılmadığı varsayalım. Bu durumda, yığın ortalaması μ 'nün tahmini yalnızca N_1 yanıtlayanlar tabakasından seçilen n_1 örneğinden tahmin ediliyor demektir (33). Bu durumda örnek ortalamasının beklenen değeri yanıtlayanların yığın ortalamasına eşit olur.

$$E(\bar{x}) = \mu_1$$

$$\begin{aligned} \bar{x} & : \text{Örnek ortalaması} \\ \bar{x}_1 & : \text{Yanıtlayanların örnek ortalaması} \\ \bar{x}_2 & : \text{Yanıtlamayanların örnek ortalaması} \end{aligned}$$

Dolayısıyla, yanıtlamama nedeniyle yanlılık aşağıdaki gibi bulunur.;

$$\begin{aligned} B(\bar{x}) & = \mu_1 - \mu \\ & = \mu_1 - (N_1/N) \cdot \mu_1 + (N_2/N) \cdot \mu_2 \\ & = (N_2/N) \cdot (\mu_1 - \mu_2) \end{aligned} \quad [1]$$

Yukarıda verilen [1] formülünden yanıtlamama yanının, N_1 yanıtlayıcı tabakasından bağımsız olduğu, ancak yanıtlamama oranına bağlı olduğu, yanıtlamama oranı arttıkça bu yanın artacağı görülmektedir. Buradan, yanıtlamama yanının, örnek çapının arttırılmasıyla azaltılamayacağı, farklı yöntemlerin uygulanması gerektiği anlaşılmaktadır.

Anket formunda yer alan sorular mümkün olan en yüksek yanıt oranına ulaşılacak şekilde tasarlanmalıdır. Soruları yanıtlayanların, yanıt vermeyenlerden farklı olup olmadıkları incelenmelidir. Fakat yanıtlamayanlar hakkında hiçbir şey bilinmiyorsa, yanıtlamama yanısı düzeltilmesi için hiçbir şey yapılamaz (33).

3. YANITLAMAMA İÇİN ÇİFT ÖRNEKLEME YÖNTEMİ

Örnek birimleri bazı arařtırmalarda bir ařamada örneęe seçilir. Bu yönleme tek ařamalı örnekleme yöntemi adı verilir. Eęer örnek birimleri birden çok ařamada örneęe seçiliyor ise, ařama sayısına göre iki ařamalı örnekleme, üç ařamalı örnekleme,... yöntemi adlarını alır. Bu yönleme genel olarak çok ařamalı örnekleme yöntemi adı verilir. Bazı arařtırmalarda, birinci ařamada alınan örnek birimlerinin içerdikleri yığın birimleri benzer özellik gösterirler. Bu durumda örnek birimlerinin tümünü incelemek yerine bundan yine bir örnek seçmek para, zaman ve emek yönünden tasarruf sağlayacağı gibi sonuçta örneğin yığını temsil etme niteliğini etkilemez (37).

İki ařamalı örnekleme veya dięer adıyla çift örnekleme, N genişlikli yığından tesadüfi olarak seçilen n çaplı örnekten, n' çaplı ($n' < n$) bir tesadüfi altörnek seçiminin yapılması işlemidir (Neyman, 1938). Böylece örnekleme alınacak yığın birimlerine iki ařamada ulařılır. Birinci ařamada seçilen örnek birimlerine ilk örnek birimleri (primary unit), ikinci ařamada her ilk örnekten seçilen yığın birimlerine altörnek birimler adı verilir (37).

Çift örnekleme yönteminin yanıtlanama durumuna uygulanması, tabaklama için çift örnekleme yönteminin özel bir durumudur (1). Tabaklama için çift örnekleme yönteminde, N birimden oluşan yığından tesadüfi olarak seçilen n çaplı örnek, h. tabakada n_h ($h=1, \dots, L$) birim olacak şekilde tabakalara ayrılır. Daha sonra tabakalardan bir altörnek seçimi yapılır ve altörnekte bulunan her bir birim için ilgilenilen deęişken deęeri elde edilir (38, 39).

Yanıtlanama olması durumu için çift örnekleme yönteminde, yanıtlayanlar ve yanıtlanmayanlar iki ayrı tabaka olarak sınıflandırılır. Birim örneęe seçilirken yanıtlanmayan olup olmadığı bilinmez; ancak örneęe seçildikten sonra gözlem sonuçlarından ortaya çıkar (38). Bu yöntemde, yanıtlanmayanlar ayrı bir tabaka olarak sınıflandırılır ve yanıtlanmayanlar tabakasından bir tesadüfi altörneęi seçilir. Seçilen

altörneğe, ilk örnektekine kıyasla daha yoğun çabanın harcandığı bir yöntem veya daha farklı bir veri toplama tekniği kullanılarak yanıt alınmaya çalışılır.

Yanıtlamama için çift örnekleme yöntemi, yanıtlamayanların tümüyle yapılacak yeniden görüşmelerin çok pahalı olması durumunda ekonomik olacaktır (40). Yanıtlamama olması durumu için üç veya daha fazla aşama örnekleri de kullanılabilir, ancak zaman ve maliyet kısıtlamaları çoğunlukla daha fazla aşama örneklerinin kullanılması işlemine izin vermez (41).

Bu yöntemle yanıtlamayanlar tabakasından altörnek seçimi tesadüfi olarak yapıldığı için, tesadüfi altörnek birimlerinin tümünden yanıt alınması durumunda, yanıtlamama yanlılığı giderilmiş olur (1).

Yanıtlamama yanlılığının giderilmesi ve yanıt oranının artırılması için önerilen yöntemlerden biri olan çift örnekleme yöntemi, Hansen ve Hurwitz (9), Cochran (10), Srinath (11) ve Deming (12) tarafından ele alınmıştır. Araştırmamız kapsamında bu yöntemlerin tümüne yer verilmiştir.

3.1. Hansen ve Hurwitz Yöntemi

Hansen ve Hurwitz (9), yanıtlamama olması durumunda, araştırma değişkenleri açısından genellikle yanıtlamayanların yanıtlayanlardan farklı özelliklere sahip olduğunu, bu nedenle yalnızca yanıtlayanlar verisine dayalı yapılacak tahminlerin yanlı olacağını savunmuşlardır. Eğer büyük çaplı yanıtlamama oranı var ise, durumun çok daha ciddi boyutta olacağını belirtmişlerdir.

Bu nedenle yanıtlamama yanını gidermek veya en aza indirmek için, yanıtlamayan birimlerinden yanıt alınmak üzere çeşitli girişimlerin, yeniden görüşmelerin yapılması gerekir. Ancak tüm yanıtlamayan birimlerinden yanıt almaya çalışılmak çok yüksek maliyetli olacağından, yanıtlamama oranının yüksek olduğu durumlarda yanıtlamayanlardan altörnek seçimi yapılarak, yalnızca altörnek birimlerinden yanıt alınmaya çalışmak daha az maliyetli olacaktır.

Her bir anket yönteminin farklı yanıt oranı ve farklı maliyetleri vardır. Hansen ve Hurwitz (9), posta anketlerinde yanıtlanama sorununu ele almıştır. Birçok araştırmada posta anketi kullanılmaktadır. Bunun nedeni, diğer veri toplama yöntemlerine kıyasla daha düşük maliyetli olmasıdır. Ancak, posta anketi yönteminde, adreslere postalanan anket formlarının geri dönüşümlerinin düşük olduğu, bunun sonucu olarakta büyük çaplı yanıtlanama oranlarının ortaya çıktığı gözlenmektedir (1).

Posta anketi yöntemi 1920'li ve 30'lu yıllarda yoğun olarak kullanılmıştır ve hâlâ bir çok alan araştırmasında düşük maliyetli ve kolay uygulanır olması nedeniyle yaygın bir şekilde kullanılmaktadır. Ancak, posta anketi yöntemine temel itiraz, yanıt vermemenin sebep olduğu sapma ve yanıt vermeyenlerin yüzdesinin büyük olmasıdır (20).

Yüz-yüze görüşme anketi yöntemiyle genellikle birimlerin hemen hemen tamamından yanıt alınabilmektedir, ancak birim başına görüşme maliyeti diğer yöntemlere kıyasla, özellikle posta anketi yöntemine kıyasla çok yüksektir (19). Hansen ve Hurwitz (9), yanıtlanama olması durumu için posta ve yüz-yüze görüşme anketi yöntemlerinin her ikisinin avantajlarından en iyi şekilde faydalanılacak bir çift örnekleme yöntemi önermiştir.

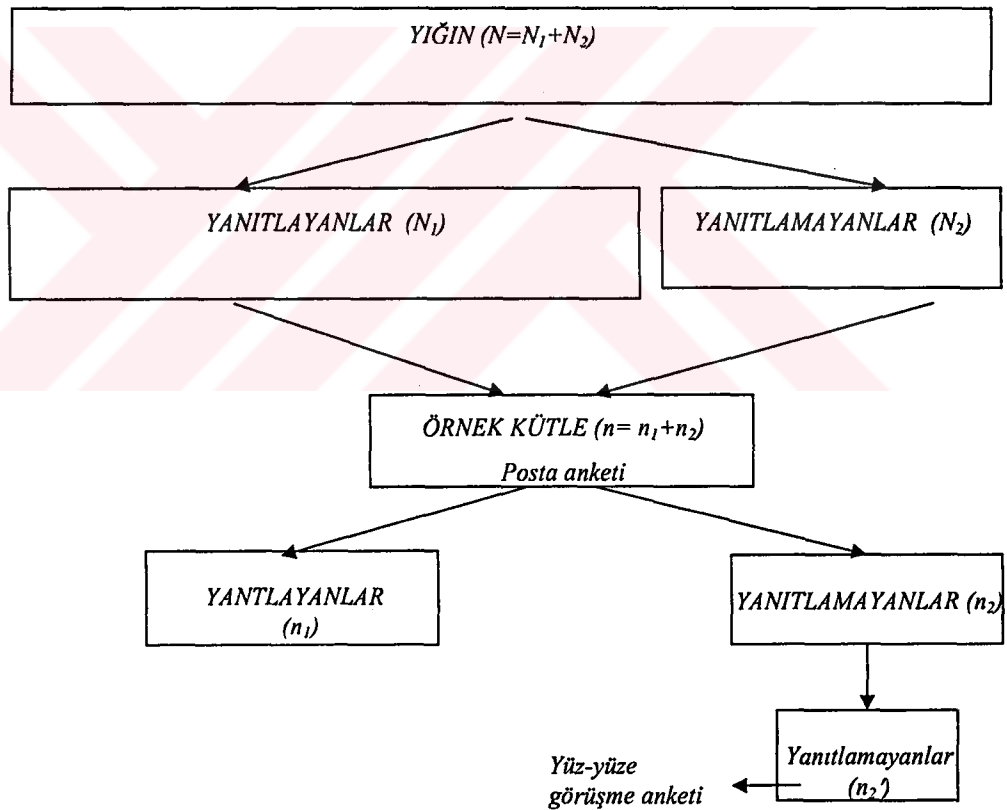
Yığından tesadüfi olarak seçilen n çaplı örnek birimlerinden n_1 tanesinden yanıt alındığı, $n_2 = n - n_1$ tanesinden yanıt alınmadığı varsayılır. Yanıt alınamayan n_2 birim arasından, $n_2' = n_2 / k$ (k : altörnekleme kesri olup, 1'den büyük sabit bir değere eşittir) çaplı bir altörnek seçimi yapılarak, seçilen altörnek birimlerinden veri toplanmaya çalışılır.

Yanıtlanama olması durumunda istenilen özelliklere ilişkin yansız tahminlerin çift örnekleme yöntemiyle yapılması süreci aşağıdaki gibi verilebilir:

1. Yığından tesadüfi seçilen n çaplı örnek birimlerinin tamamına düzenlenen anket formları postalanır,

2. Posta anketleri için belirlenmiş olan geri dönüşüm süresinin sona ermesiyle, anket formlarını geri gönderenler (yanıtlayanlar) ve göndermeyenler (yanıtlamayanlar) tespit edilir,
3. Posta anketlerini yanıtlamayanlar örneğinden tesadüfi seçim yöntemiyle bir altörnek alınır. Seçilen altörnek birimleriyle yüz-yüze görüşme yapılır,
4. Bu süreçlerin sonunda, yığma ilişkin istenilen özelliklerin tahminlerini yapmak üzere posta anketlerini yanıtlayanlar örneğinden ve altörnekten toplanan veriler birleştirilir.

Ayrıca bu süreci, şematik olarak aşağıdaki gibi göstermek de mümkündür.



Şekil 3. 1. Yanıtlamama olması durumunda çift örnekleme yöntemi

Araştırma konusu birimlere posta anketi gönderilmesi yüz-yüze görüşme anketi uygulamasında daha ekonomik olacaksa, ilk aşamada örnek birimlerine anketlerin postalanması yerinde bir girişim olacaktır (6).

3.1.1. Yığın ortalamasının tahmini ve tahminin varyansı

Çift örnekleme yöntemiyle, yığına ilişkin ortalama ve toplam parametrelerinin yansız tahminleri ve tahminlere ilişkin varyansları elde edilmeye çalışılmıştır. Bu yöntemle \bar{X} yığın ortalaması ve X yığın toplamı için yansız bir tahmin şöyle verilebilir.

$$\bar{x}' = \frac{n_1}{n} \bar{x}_1 + \frac{n_2}{n} \bar{x}_2' \quad [2]$$

\bar{x}' : Yığın ortalamasının yansız bir tahmini

\bar{x}_1 : n_1 yanıtlayanlar tabakasının ortalama değeri

\bar{x}_2' : Altörnek ortalama değeri

n_1 : Örnek birimlerinden yanıtlayanların sayısı

n_2 : Örnek birimlerinden yanıtlamayanların sayısı

n : Örnek çapı

$$\hat{X}' = N \cdot \bar{x}' \quad [3]$$

\hat{X}' : Yığın toplamının yansız bir tahmini

\bar{x}' 'nin yığın ortalamasının yansız bir tahmin edicisi olduğu aşağıdaki gibi verilmiştir.

İSPAT:

$E(\bar{x}') = \bar{X}$ ise, \bar{x}' yansız bir tahmin edicidir denir. \bar{x}' 'nin, \bar{X} yığın ortalamasının yansız bir tahmini olduğu aşağıdaki gibi gösterilmiştir.

$$E(\bar{x}') = E\left(\frac{n_1}{n} \bar{x}_1 + \frac{n_2}{n} \bar{x}_2'\right)$$

$$= E\left(\frac{n_1}{n}\bar{x}_1\right) + E\left(\frac{n_2}{n}\bar{x}_2\right) \quad [4]$$

[20] eşitliğinin sağ tarafındaki birinci terim ele alınacak olursa,

$$\begin{aligned} E\left(\frac{n_1}{n}\bar{x}_1\right) &= E\left[\frac{n_1}{n}E(\bar{x}_1)\right] \\ &= \frac{N_1}{N}\bar{X}_1 \end{aligned} \quad [5]$$

olarak bulunur.

Yanıtlamama için çift örnekleme yöntemi, yanıtlamayanlardan tesadüfilik varsayımı altında bir altörnek seçim işlemidir. Dolayısıyla, yanıtlamayanların tamamı ölçülmemiş olduğu halde, tesadüfilik varsayımı altındaki çift örnekleme işlemi yanıtlamamadan kaynaklanan bir yanı içermez (37, 1). Bu nedenle, $E(\bar{x}_2') = \bar{x}_2$ ve $E(\bar{x}_2) = \bar{X}_2$ dir. Buradan \bar{x}_2' 'nin yansız bir tahmin olduğu aşağıdaki gibi elde edilmiş olacaktır.

$$\begin{aligned} E\left(\frac{n_2}{n}\bar{x}_2'\right) &= E\left[\frac{n_2}{n}E(\bar{x}_2')\right] \\ &= E\left[\frac{n_2}{n}\bar{x}_2\right] \\ &= \frac{N_2}{N}\bar{X}_2 \end{aligned} \quad [6]$$

[4] denkleminde, [5] ve [6] sonuçları yerine yazılırsa, \bar{x} 'nin yığın ortalamasının yansız bir tahmini olduğu gösterilmiş olur.

$$E(\bar{x}) = \frac{N_1}{N}\bar{X}_1 + \frac{N_2}{N}\bar{X}_2$$

$$= W_1 \bar{X}_1 + W_2 \bar{X}_2$$

$$E(\bar{x}) = \bar{X}$$

Çift örnekleme yöntemi ile \hat{X}' 'nin varyansı,

$$V(\hat{X}') = N^2 \cdot \frac{N-n}{(N-1)} \frac{\sigma^2}{n} + \frac{N}{n} (k-1) \frac{N_2}{N_2-1} \sigma_2^2 \quad [7]$$

- N : Yığın çapı
 σ^2 : Yığın varyansı
 σ_2^2 : Yanıtlamayanlar tabakasının varyansı
 $k=n_2/n_2'$: Altörnekleme kesri

şeklinde verilmiştir. Bu varyans formülüne nasıl ulaşıldığının ispatı aşağıdaki gibidir.

İSPAT:

$V(\hat{X}') = E(\hat{X}' - X)^2$ olarak tanımlanır. Buradan,

$$V(\hat{X}') = E \left\{ \frac{N}{n} (n_1 \bar{x}_1 + n_2 \bar{x}_2') - X \right\}^2$$

$$= N^2 E \left\{ \frac{(n_1 \bar{x}_1 + n_2 \bar{x}_2')}{n} - \bar{X} \right\}^2$$

olur.

Yığın ortalaması yanıtlayanlar ve yanıtlamayanlar tabakasına dayalı olarak aşağıdaki şekilde tanımlanırsa;

$$\bar{x} = \frac{n_1}{n} \bar{x}_1 + \frac{n_2}{n} \bar{x}_2$$

$n_1 \bar{x}_1$ aşağıdaki gibi yazılabilir.

$$n_1 \bar{x}_1 = n\bar{x} - n_2 \bar{x}_2$$

olarak alınır ve varyans formülünde yerine yazılırsa, varyans formülü şöyle yazılabilir:

$$\begin{aligned} V(\hat{X}') &= \frac{N^2}{n^2} E \left[n\bar{x} - n_2 \bar{x}_2 + n_2 \bar{x}_2' - n\bar{X} \right]^2 \\ &= \frac{N^2}{n^2} E \left[n(\bar{x} - \bar{X}) + n_2(\bar{x}_2' - \bar{x}_2) \right]^2 \\ &= \frac{N^2}{n^2} E \left\{ E \left[n^2 (\bar{x} - \bar{X})^2 \right] + 2E \left[n(\bar{x} - \bar{X})n_2(\bar{x}_2' - \bar{x}_2) \right] + E \left[n_2^2 (\bar{x}_2' - \bar{x}_2)^2 \right] \right\} \quad [8] \end{aligned}$$

n gözlemlerinin sabit bir kümesi için, \bar{x} , n_2 , \hat{X} ve \bar{x}_2 değerleri de sabittir. Dolayısıyla, yanıtlamayanlar altörneğinden yapılan tahminlerde yansızdır. $E(\bar{x}_2') = \bar{x}_2$ olduğundan, [8] eşitliğinde ortadaki terim sıfır olur. Yanıtlamayan birimlerin tüm mümkün altkümeleri üzerinden beklenen değer alındığında, üçüncü terim;

$$\frac{N^2}{n^2} E \left[n_2^2 (\bar{x}_2' - \bar{x}_2)^2 \right] = \frac{N^2}{n^2} n_2^2 \left[\frac{n_2 - n_2'}{n_2} \frac{\sum_{i=1}^{n_2} (x_i - \bar{x}_2)^2}{(n_2 - 1)n_2'} \right]$$

şeklinde yazılabilir. n_2 yanıtlamayan birimlerinin tüm mümkün altkümeleri üzerinden beklenen değer alındığında, üçüncü terim aşağıdaki gibi bulunmuş olacaktır.

$$\frac{N^2}{n^2} n_2^2 \frac{n_2 - n_2'}{(n_2 - 1)n_2} \frac{N_2}{n_2} \frac{n_2 - 1}{N_2 - 1} \sigma_2^2 \quad [9]$$

k altörnekleme kesri, n_2' altörnek çapı olmak üzere, $n_2' = n_2/k$ dan, [9] eşitliğini k 'ya bağlı olarak aşağıdaki gibi yazmak da mümkündür.

$$\frac{N^2}{n^2} n_2 \frac{N_2}{N_2 - 1} (k-1) \sigma_2^2 \quad [10]$$

N_2 : Yığındaki yanıtlayanların sayısı

σ^2 : Yığın varyansı

σ_2^2 : N_2 yanıtlayanların varyansı

n_2 örnekten örneğe farklılık gösterdiğinde, tüm mümkün örnek durumları için [10]'nun beklenen değerini alınmaktadır. Bu durumda üçüncü terim,

$$\frac{N^2}{n^2} N_2 \frac{N_2}{N_2 - 1} (k-1) \frac{n}{N} \sigma_2^2 = \frac{N}{n} \frac{N_2^2}{N_2 - 1} (k-1) \sigma_2^2 \quad [11]$$

olarak elde edilir.

[8] eşitliğinde birinci terimi,

$$\frac{N^2}{n^2} E[n^2 (\bar{x}' - \bar{X})^2] = N^2 E[(\bar{x}' - \bar{X})^2]$$

yerine koymadan örneklemede $E[(\bar{x}' - \bar{X})^2]$ aşağıdaki gibi tanımlandığında,

$$E[(\bar{x}' - \bar{X})^2] = \frac{N-n}{(N-1)} \frac{\sigma^2}{N}$$

beşinci terim aşağıdaki şekilde olacaktır.

$$\frac{N^2}{n^2} E[n^2 (\bar{x}' - \bar{X})^2] = N^2 \frac{N-n}{(N-1)n} \sigma^2 \quad [12]$$

σ^2 : Yığın varyansı

Bulunan [11] ve [12] sonuçları, [8] eşitliğinde yerine yazıldığında, çift örnekleme yöntemi ile \hat{X} için varyans formülüne aşağıdaki gibi ulaşılmış olur.

$$V(\hat{X}') = N^2 \cdot \frac{N-n}{(N-1)n} \sigma^2 + \frac{N}{n} (k-1) \frac{N_2}{N_2-1} \sigma_2^2 \quad [7]$$

Yığın toplamı için verilen varyans formülünde, N ve N₂ yeteri kadar geniş ve $\sigma^2 = \sigma_2^2$ iken, $\frac{N}{N-1}$ ve $\frac{N_2}{N_2-1}$ 'nün yaklaşık olarak 1'e eşit olduğu varsayımı

yapılmış olsun. Ayrıca, yığından tesadüfi olarak seçilen \hat{n} çaplı örneğe posta anketi yöntemi uygulanmış ve tüm örnek birimlerinin yüzde 100'ünden yanıt alınmış ise, yığın toplamı için varyans formülü aşağıdaki gibi olacaktır.

$$V(\hat{X}') = N^2 \cdot \frac{N-\hat{n}}{(N-1)\hat{n}} \sigma^2 \quad [13]$$

σ^2 : Yığın varyansı

N : Yığın çapı

\hat{n} : Posta anketi gönderilen birimlerin sayısı

3.1.3. n ve k'nın optimum değerleri

Hansen ve Hurwitz, posta anketlerini yanıtlamayanlara tekrar anket formları postalamak için harcanacak parayı, yüz-yüze görüşme yapılmak üzere yanıtlamayanlardan seçilen altörnek birimlerine harcamak daha faydalı olacağını belirtmiştir. Araştırmanın toplam maliyetini minimum yapan optimum ilk örnek genişliği n ve optimum altörnekleme kesri k değerleri bulunmaya çalışılacaktır.

Çift örnekleme yöntemiyle örnekleme yapma maliyeti aşağıda verilen maliyet fonksiyonuyla tanımlanmıştır.

$$C = c_0n + c_1n_1 + c_2n_2' \quad [14]$$

- C : Örnekleme yapma maliyeti
 c_0 : Birim başına anket formu postalama maliyeti
 c_1 : Yanıtlanan bir posta anketinin işlenilmesi maliyeti
 c_2 : Birim başına yüz-yüze görüşme yapılması ve alınan sonuçların işlenilmesi maliyeti

Yukarıda verilen maliyet fonksiyonunu aşağıdaki gibi tanımlamakta mümkündür.

$$C = c_0n + c_1pn + c_2 \frac{qn}{k}$$

$$p = \frac{n_1}{n}, \quad q = 1 - p = \frac{n_2}{n}$$

p : Posta anketini yanıtlayanların oranı

q : Posta anketini yanıtlamayanların oranı

\bar{x}' 'nin varyansı $V(\bar{x}') = \frac{V(\hat{X})}{N^2}$ eşitliğinden aşağıdaki gibidir.

$$V = V(\bar{x}') = \frac{N-n}{(N-1)n} \sigma^2 + \frac{1}{n.N} (k-1) \frac{N_2}{N_2-1} \sigma_2^2$$

n ve k 'nin optimum değerleri, $V=V(\bar{x}')$ varyans ve C maliyet kısıtları altında,

$C(V + \frac{\sigma^2}{N-1})$ fonksiyonunu minimum yapılacak şekilde seçilir.

k 'nin optimum değerini hesaplayabilmek için, $C(V + \frac{\sigma^2}{N-1})$ fonksiyonunun k 'ya göre türevi alınıp sifıra eşitlenir. Türevi alınıp sifıra eşitlenen bu fonksiyonda k yalnız

bırakılacak olursa, k'nın optimum değeri elde edilmiş olacaktır. Optimum k'nın hesaplanması süreci aşağıdaki gibi gösterilmiştir.

$$\begin{aligned}
V + \frac{\sigma^2}{N-1} &= \frac{N-n}{(N-1)n} \sigma^2 + \frac{1}{n.N} (k-1) \frac{N_2}{N_2-1} \sigma^2 + \frac{\sigma^2}{N-1} \\
&= \frac{N\sigma^2}{(N-1)n} - \frac{\sigma^2}{N-1} + \frac{kW_2\sigma_2^2}{n.(N_2-1)} - \frac{W_2\sigma_2^2}{n.(N_2-1)} + \frac{\sigma^2}{N-1} \\
&= \frac{N\sigma^2}{(N-1)n} + \frac{kW_2\sigma_2^2}{n.(N_2-1)} - \frac{W_2\sigma_2^2}{n.(N_2-1)} \\
&= \frac{N(N_2-1)\sigma^2 - (N-1)W_2\sigma_2^2}{(N-1)(N_2-1)n} + \frac{kW_2\sigma_2^2}{n.(N_2-1)} \\
C(V + \frac{\sigma^2}{N-1}) &= (c_0n + c_1pn + c_2 \frac{qn}{k}) \left(\frac{N(N_2-1)\sigma^2 - (N-1)W_2\sigma_2^2}{(N-1)(N_2-1)n} + \frac{kW_2\sigma_2^2}{n.(N_2-1)} \right) \\
&= \frac{(c_0n + c_1pn)N(N_2-1)\sigma^2}{(N-1)(N_2-1)n} - \frac{(c_0n + c_1pn)(N-1)W_2\sigma_2^2}{(N-1)(N_2-1)n} \\
&\quad + \frac{c_2qnN(N_2-1)\sigma^2}{k(N-1)(N_2-1)n} - \frac{c_2qn(N-1)W_2\sigma_2^2}{k(N-1)(N_2-1)n} + \frac{(c_0n + c_1pn)kW_2\sigma_2^2}{n.(N_2-1)} + \frac{c_2qnkW_2\sigma_2^2}{kn.(N_2-1)}
\end{aligned}$$

Bu denklemin k'ya göre türevi alınıp sıfıra eşitlenirse, optimum k değeri hesaplanmış

olur. $\frac{\partial(V + \frac{\sigma^2}{N-1})C}{\partial k} = 0$ denklemini çözümü, optimum k'yı verir.

$$-\frac{c_2qnN(N_2-1)\sigma^2}{k_{opt}^2(N-1)(N_2-1)n} + \frac{c_2qn(N-1)W_2\sigma_2^2}{k_{opt}^2(N-1)(N_2-1)n} + \frac{(c_0n + c_1pn)W_2\sigma_2^2}{n.(N_2-1)} = 0$$

$$k_{opt}^2 = \frac{c_2 q N (N_2 - 1) \sigma^2 - c_2 q (N - 1) W_2 \sigma_2^2}{(N - 1)(c_0 + c_1 p) W_2 \sigma_2^2}$$

$$k_{opt} = \sqrt{\frac{c_2 q N (N_2 - 1) \sigma^2 - c_2 q (N - 1) W_2 \sigma_2^2}{(N - 1)(c_0 + c_1 p) W_2 \sigma_2^2}} \quad [15]$$

verilir.

Yanıtlamama olması durumunda yığından alınacak optimum örnek çapı,

$$n_{opt} = \hat{n} \left\{ 1 + (k - 1) q \frac{N - 1}{N_2 - 1} \frac{\sigma^2}{\sigma_2^2} \right\}, \quad [16]$$

\hat{n} : Posta anketi gönderilmek üzere seçilen birim sayısı

şeklindedir. Optimum altörnekleme kesri,

$$n_2' = \frac{n_2}{k_{opt}}$$

hesaplanır. p , $q = 1 - p$ değerleri ile, $\sigma^2 = \sigma_2^2$ ve $\frac{N}{N - 1} = 1, \frac{N_2}{N_2 - 1} = 1$ varsayımı altında optimum n , k ve n_2' değerleri sırasıyla şöyle belirlenir:

$$n_{opt} = \hat{n} \{ 1 + (k - 1) q \}, \quad q = 1 - p \quad [17]$$

[15] denklemiyle verilen k_{opt} formülünde $P, Q=1-P$ değerleri ile, $\sigma^2 = \sigma_2^2$ ve

$$\frac{N}{N-1} = 1, \frac{N_2}{N_2-1} = 1 \quad (N = N-1, N_2 = N_2-1) \text{ varsayımı altında optimum } k;$$

$$k_{opt} = \sqrt{\frac{c_2 \frac{N_2}{N} N N_2 \sigma^2 - c_2 \frac{N_2}{N} N \frac{N_2}{N} \sigma^2}{N(c_0 + c_1 P) \frac{N_2}{N} \sigma^2}}$$

$$k_{opt} = \sqrt{\frac{c_2 - c_2 N_2}{c_0 + c_1 P}}, \quad P = 1 - N_2 \text{ olduğu için optimum altörnekleme kesri}$$

aşağıdaki gibi bulunmuş olur.

$$k_{opt} = \sqrt{\frac{c_2 P}{c_0 + c_1 P}}$$

[18]

$$n_2' = n_2 / k$$

3.1.3. Oran ve regresyon tahmini

Yığın ortalaması tahmini yapılacak X_i değişkenine ilişkin bir Y_i yardımcı değişken bilgisi mevcut ise, yığın ortalaması veya toplamı tahmini oran veya regresyon tahmini ile yapılabilir. Yanıtlamama için çift örnekleme yönteminde, ortalamanın oran tahmini, bir önceki kesimde [2] formülüyle verilen ortalama tahmininden daha duyarlı sonuçlar verir.

Çift örnekleme yöntemiyle yığın ortalamasının oran tahmini aşağıdaki şekilde tanımlanmıştır:

$$\bar{x}_0' = \frac{n_1 \bar{x}_1 + n_2 \bar{x}_2}{n_1 \bar{y}_1 + n_2 \bar{y}_2} \bar{y} \quad [19]$$

\bar{Y} : Daha önceki kayıtlardan yığının ortalama değeri

\bar{y}_1 : Daha önceki kayıtlardan posta anketi yanıtlayanlara ilişkin ortalama değeri

\bar{y}_2 : Daha önceki kayıtlardan yüz-yüze görüşme anketi yapılanlardan elde edilen ortalama değeri

Yığın toplamı için oran tahmini de aşağıdaki gibi olacaktır.

$$\hat{X}_{o'} = N \cdot \bar{x}_{o'}$$

$\hat{X}_{o'}$ tahmini için varyans yaklaşımı, [7] ile verilen $V(\hat{X}_{o'})$ varyans formülünden hesaplanır. Böylece varyans formülü,

$$V(\hat{X}_{o'}) = N^2 \cdot \frac{N-n}{(N-1)n} \sigma_1^2 + \frac{N}{n} (k-1) \frac{N_2}{N_2-1} \sigma_2^2 \quad [20]$$

şeklinde olur. [20] formülde verilen S_1^2 aşağıdaki gibi tanımlanmıştır:

$$\sigma_1^2 = \sigma_x^2 - 2R\sigma_{xy} + \sigma_y^2$$

σ_2^2 ise aşağıdaki gibi tanımlanmıştır.

$$\sigma_2^2 = \sigma_{2x}^2 - 2R\sigma_{2(xy)} + \sigma_{2y}^2$$

Burada σ_x^2 , [7] formülünde verilen σ^2 değeri ile aynıdır. σ_y^2 ; yardımcı değişkenin varyans değeri, σ_{xy} ; x ile y arasındaki kovaryans ve $R = \frac{\bar{x}}{\bar{y}}$ dir.

Optimum n ve n_2' , [4] formülünde, σ^2 ve σ_2^2 değerleri yerine yukarıda tanımlanmış değerlerin yerleştirilmesiyle [20] formülüyle bulunur. Böylece optimum n ve n_2' değerleri [20] ve [11]formüllerinden elde edilebilir.

3.2. Cochran'ın Yöntemi

Cochran (10), çift örnekleme yöntemini Hansen ve Hurwitz (9)'in yöntemine dayanarak, "Yanıtlamayanlar Arasındaki Optimum Örnekleme Kesri" başlığı altında incelemiştir. Çift örnekleme yönteminin tabakalama için çift örnekleme yönteminin bir uygulaması olarak ele alınabileceği belirtilmiştir. Tabakalama için çift örnekleme yöntemi çerçevesinde yığın, ilk girişimde yanıtlayanlar ve yanıtlamayanlar olarak iki tabakaya ayrılmış olduğu, yığından seçilen n çaplı örneğe postalanan anket formlarının n_1 tanesinden yanıt alındığı, $n_2 = n - n_1$ tanesinden yanıt alınmadığı varsayımı yapılmıştır. n_2 yanıtlamayanlarından $n_2' = n_2 / k$ çaplı, ($k \geq 1$) bir tesadüfi altörneği seçim işleminin yapılması önerilmiştir.

3.2.1. Yığın ortalamasının tahmini ve tahminin varyansı

\bar{X} yığın ortalaması ve X yığın toplamının yansız bir tahmini, Hansen ve Hurwitz'in yönteminde olduğu gibi aşağıdaki gibi verilmiştir:

$$\bar{x}' = \frac{n_1}{n} \bar{x}_1 + \frac{n_2}{n} \bar{x}_2' \quad [21]$$

\bar{x} : Yığın ortalamasının yansız bir tahmini
 \bar{x}_1 : n_1 yanıtlayanlar tabakasının ortalama değeri
 \bar{x}_2' : Altörneğin ortalama değeri

$$\hat{X}' = N \cdot \bar{x}' \quad [22]$$

\hat{X} : Yığın toplamının yansız bir tahmini

Rao (42), yanıtlamama olmadığı durum için yığın varyansını,

$$V(\bar{x}/s^2) = \frac{N-n}{N} \frac{S^2}{n}$$

olarak verirken, s_2^2 yanıtlamayanlar tabakasının örnek varyansının tahminini,

$$v(\bar{x}/s_2^2) = w_2 \frac{(k-1) S_2^2}{n_2}$$

$$w_2 = n_2/n$$

şeklinde olduğunu göstermiştir.

$\frac{n_2}{n} = \frac{N_2}{N}$ olduğunda, Cochran'da, çift örnekleme yöntemiyle \bar{x}' 'nin varyansı aşağıdaki şekilde verilmektedir.

$$V(\bar{x}') = \frac{N-n}{N} \frac{S^2}{n} + \frac{(k-1) W_2 S_2^2}{n} \quad [23]$$

[23] eşitliğinin sağ tarafında yer alan ikinci terim, altörneklemeden dolayı varyansta oluşan artıştır (42). Bu artış, W_2 yanıtlamama oranı ve S_2^2 varyansı düşük, $(1/k)$ altörnekleme kesri büyük olduğunda düşük olacaktır. Seçilen n örneğinde yanıtlamama durumu olmasaydı, $V(\bar{x}')$ varyans formülünün ikinci terimi olmayacak ve yapılan tahminler daha duyarlı olacaktır.

Cochran (10) ve Rao (42) çıkarsamalarından, varyans tahmini aşağıdaki şekilde tanımlanmıştır.

$$v(\bar{x}') = \frac{(N-n)(n_1-1)}{N(n-1)} w_1 \frac{s_1^2}{n_1} + \frac{(N-1)(n_2-1) - (n-1)(n_1-1)}{N(n-1)} w_2 \frac{s_2^2}{n_2}$$

$$+ \frac{N-n}{N(n-1)} \left[w_1 (\bar{x}'_1 - \bar{x}'_2)^2 + w_2 (\bar{x}'_2 - \bar{x}'_1)^2 \right] \quad [24]$$

s_1^2 : n_1 birimlerinin varyansı

s_2^2 : n_2 birimlerinin varyansı

3.2.2. n , n_2' ve k 'nin optimum değerleri

[23] formülünde, $k = n_2/n_2'$ mevcut koşullar dikkate alınarak veya keyfi olarak sabit kabul edilebilir. Ancak, maliyet ve yığın varyansını minimum yapacak n ve k 'nin optimum değerlerini elde etmek daha rasyonel bir yaklaşım olacaktır. Aşağıda, belirlenen maliyet fonksiyonu altında, optimum n , k ve n_2' değerleri hesaplanmaya çalışılmıştır.

Örnekleme maliyetinin, [14]denkleminde aşağıdaki gibi olduğu verilmişti.

$$C = c_0n + c_1n_1 + c_2n_2'$$

Bu maliyet fonksiyonunu aşağıdaki gibi tanımlamakta mümkündür.

$$C = c_0 + c_1n_1 + c_2 \frac{n_2}{k} \quad [25]$$

Formül [23]'le verilen varyans eşitliğinde, V varyans eşitliğinin aşağıdaki gibi olduğu gösterilmiştir.

$$V = V(\bar{x}') = \left(\frac{N-n}{N} \frac{S^2}{n} \right) + \frac{(k-1) W_2 S_2^2}{n}$$

n ve k 'nin optimum değerleri, V varyans ve C maliyet kısıtları altında, $C(V + \frac{1}{N} S^2)$ fonksiyonunu minimum yapılacak şekilde seçilir.

k 'nin optimum değerini hesaplayabilmek için, $C(V + \frac{1}{N} S^2)$ fonksiyonunun k 'ya göre türevi alınıp sifıra eşitlenir. Türevi alınıp sifıra eşitlenen bu fonksiyonda k yalnız bırakılacak olursa, k 'nin optimum değeri elde edilmiş olacaktır. Optimum k 'nin hesaplanması süreci aşağıdaki gibi gösterilmiştir.

$V + \frac{1}{N} S^2$ denkleminde, [23]'teki V varyans formülü yerine yazılacak olursa,

$$\begin{aligned} V + \frac{1}{N} S^2 &= \frac{1}{n} S^2 - \frac{1}{N} S^2 + \frac{kW_2 S_2^2}{n} - \frac{W_2 S_2^2}{n} + \frac{1}{N} S^2 \\ &= \frac{S^2 - W_2 S_2^2}{n} + \frac{kW_2 S_2^2}{n} \end{aligned} \quad [26]$$

ve [25]'teki C maliyet fonksiyonu k'ya bağlı olarak aşağıdaki gibi tanımlanırsa,

$$C = (c_0 + c_1 W_1) n + \frac{c_2 W_2 n}{k} \quad [27]$$

denkleminde ulaşılmış olunur. Buradan, $C(V + \frac{1}{N} S^2)$ fonksiyonunda, [26] ve [27] formülleri yerine yazıldığında,

$$\begin{aligned} (V + \frac{1}{N} S^2) C &= \left[\frac{S^2 - W_2 S_2^2}{n} + \frac{kW_2 S_2^2}{n} \right] \left[(c_0 + c_1 W_1) n + \frac{c_2 W_2 n}{k} \right] \\ &= \frac{S^2 - W_2 S_2^2}{n} (c_0 + c_1 W_1) n + \frac{S^2 - W_2 S_2^2}{n} \frac{(c_2 W_2 n)}{k} \\ &\quad + \frac{kW_2 S_2^2 (c_0 + c_1 W_1) n}{n} + c_2 W_2^2 S_2^2 \end{aligned}$$

denkleminde ulaşılır. Bu denklemin k'ya göre türevi alınıp sıfıra eşitlenirse, optimum

k değeri hesaplanmış olur. $\frac{\partial (V + \frac{S^2}{N}) C}{\partial k} = 0$ denklemi çözümü, optimum k'yı verir.

$$\frac{\partial (V + \frac{S^2}{N}) C}{\partial k} = 0$$

$$-\frac{(S^2 - W_2 S_2^2) c_2 W_2}{k^2} + W_2 S_2^2 (c_0 + c_1 W_1) = 0$$

$$W_2 S_2^2 (c_0 + c_1 W_1) = \frac{(S^2 - W_2 S_2^2) c_2 W_2}{k^2}$$

$$k_{opt} = \sqrt{\frac{c_2 (S^2 - W_2 S_2^2)}{S_2^2 (c_0 + c_1 W_1)}} \quad [28]$$

Buradan, yanıtlamayanlar tabakasından alınacak n_2' altörnek genişliğide aşağıdaki gibi hesaplanmış olur.

$$n_2' = \frac{n_2}{k_{opt}} \quad [29]$$

Optimum ilk örnek çapı n , [26] veya [27] eşitliklerinden hesaplanır. İlk örnek çapı, belirlenen bir V varyans değeri altında, C maliyet fonksiyonunun minimize edilmesiyle veya belirlenen bir C maliyet fonksiyonu altında, V varyans değerinin minimize edilmesiyle belirlenebilir.

Cochran, optimum n 'i, belirlenen V varyans değeri altında, C maliyet fonksiyonunun minimize edilmesiyle hesaplanacağına karar vermiştir. Buradan, [26] eşitliğinde k yerine k_{opt} yazıldığında, bu denklem aşağıdaki şekilde gösterilmiş olacaktır.

$$V + \frac{1}{N} S^2 = \frac{S^2 - W_2 S_2^2}{n} + \frac{k_{opt} W_2 S_2^2}{n}$$

Bu denklemde n yalnız bırakılacak olursa, optimum ilk örnek genişliği aşağıdaki gibi hesaplanmış olacaktır.

$$\frac{NV + S^2}{N} = \frac{S^2 - W_2 S_2^2 + k_{opt} W_2 S_2^2}{n}$$

$$\frac{NV + S^2}{N} = \frac{S^2 + (k_{opt} - 1)W_2 S_2^2}{n}$$

Buradan n yalnız bırakılacak olursa, yanıtlamama olması durumunda yığından seçilecek optimum örnek genişliği formülüne ulaşılmış olunur.

$$n_{opt} = \frac{N[S^2 + (k_{opt} - 1)W_2 S_2^2]}{NV + S^2} \quad [30]$$

3.2.3. W_2 yanıtlamayanlar oranı bilinmiyorsa

n ve k 'nın optimum değerlerinin hesaplanmasında, W_2 yanıtlamayanlar yığın oranı bilgisine ihtiyaç vardır. W_2 oranı çoğunlukla daha önceki deneyimlere dayanılarak tahmin edilmeye çalışılır. Eğer W_2 bilinmiyorsa, [23] varyans formülü için bilinen n_2' ve n değerleriyle koşullu varyans değeri kullanılır ve optimum k değeri bu koşullu varyans formülünden elde edilir (42).

$$V_{koşullu}(\bar{x}') = \frac{N - n}{N} \frac{S^2}{n} + \frac{(k - 1) n_2 S_2^2}{n^2} \quad [31]$$

Burada, [23] formülünde W_2 yerine w_2 örnek oranı kullanılmıştır.

3.2.4. Oran ve regresyon tahmin edicileri

Cochran, önsel yanıtlamama durumu var ve ilgilenilen değişken X ile yardımcı değişken Y 'nin istenilen özellikleri ölçülemiyorsa, çift örnekleme yönteminde oran ve regresyon tahminlerinin kullanımını önermiştir (42). Burada, X ve Y , n_2' yanıtlamayanlar altörneğinden elde edilir. \bar{X} için oran tahmini aşağıdaki gibidir.

$$\bar{X}_0 = \frac{\bar{x}}{\bar{y}} \cdot \bar{Y} \quad [32]$$

\bar{Y} : *Y yardımcı değişkeninin yığın ortalaması*

Burada, $\bar{x} = w_1 \bar{x}_1 + \bar{x}_2$ eşitliğinde tanımlandığı gibi, \bar{y} değeri ise $\bar{y} = w_1 \bar{y}_1 + w_2 \bar{y}_2$ dir ve (\bar{x}_1, \bar{y}_1) değerleri birinci örnekten, (\bar{x}_2, \bar{y}_2) değerleri ise ikinci örnekten hesaplanır. Oran tahminin yaklaşık bir varyansı, $d = x_i - Ry_i$ iken, [23] formülünde S^2 yerine S_d^2 , $R = \bar{X} / \bar{Y}$ ve S_2^2 yerine S_{2d}^2 ve R için $R = \bar{X}_2 / \bar{Y}_2$ değeri kullanılır. Burada, S_d^2 tüm yığın için varyans, S_{2d}^2 yanıtlamayanlar tabakası için varyans değeridir.

\bar{X} için regresyon tahmini aşağıdaki gibidir.

$$\bar{x}_{dr} = \bar{x} + b(\bar{Y} - \bar{y}) \quad [33]$$

b : Regresyon katsayısı

\bar{x}_{dr} regresyon tahminin yaklaşık bir varyansı, [22] formülünde, S^2 yerine $S^2(1 - \rho^2)$ değeri, ve S_2^2 yerine $S_2^2(1 - \rho_2^2)$ değerinin kullanılmasıyla hesaplanır. Burada, ρ^2 tüm yığın için X ve Y arasındaki ilişki katsayısını, ρ_2^2 yanıtlamayanlar tabakası için X ve Y arasındaki ilişki katsayısını ifade etmektedir.

3.3. Srinath'ın Yöntemi

El-Badry (43), Hansen ve Hurwitz (9)'in örnekleme planını posta anketini yanıtlamayanlar tabakasına tekrar posta anketleri gönderme girişimlerinin yapılması, yapılan bu girişimler sonunda hala yanıtlamayanlar mevcutsa, bu yanıtlamayanlardan yüz-yüze görüşme yapılmak üzere bir altörnek alınması olarak genişletmiştir. Yığna ilişkin tahminler bu girişimlerin ve altörnekleme sonuçlarının bir araya getirilmesiyle hesaplanmaya çalışılmıştır.

Srinath (11), altörnekler seçimi için önermiş olduğu alternatif örnekleme kuralını, Hansen ve Hurwitz (9)'in çift örnekleme planı içinde ve hem de El-Badry (43)'in çok aşamalı örnekleme planını içinde formüle etmiştir.

Sukhatme (44), Srinath'ın önermiş olduğu örnekleme planında altörnekleme kesirlerinin sabit tutulmamakla birlikte, yanıtlamama oranlarına dayalı olarak da değişme göstermemiş olduğu belirtilmiştir. Bu planda tahminin varyansı, yığındaki yanıtlamamanın bilinmeyen oranlarından bağımsız kabul edilmiştir. Bu kurallar, yanıtlamama oranları tam olarak bilinmediğinde, tahmin edicinin istenilen bir doğruluğunu elde etmek için, ilk örnek çapı ve altörnekleme kesirlerinin hesaplanabilmesine olanak sağlar.

3.3.1. Yığın ortalamasının tahmini ve tahminin varyansı

Srinath'ın çift örnekleme yöntemiyle \bar{X} yığın ortalaması tahmini, Hansen ve Hurwitz (9) formülüyle verilmiştir.

$$\bar{x}' = \frac{n_1}{n} \bar{x}_1 + \frac{n_2}{n} \bar{x}_2'$$

Buna karşın \bar{x}' için varyans formülü aşağıdaki gibi verilmiştir.

$$V(\bar{x}') = \left(\frac{N-n}{Nn}\right)S^2 + \frac{k'S_2^2}{n} \quad [35]$$

Bu varyans formülüne nasıl ulaşıldığı aşağıda verilmektedir.

İSPAT:

\bar{x}' formülünde $\frac{n_2}{n} \bar{x}_2$ terimi bir kez eklenip bir kez çıkarıldığında, ortalama formülü aşağıdaki gibi yazılmış olacaktır.

$$\bar{x}' = \frac{1}{n}(n_1\bar{x}_1 + n_2\bar{x}_2) + \frac{n_2}{n}(\bar{x}_2' - \bar{x}_2)$$

\bar{x}_2' : Yanıtlamayanların tümü için örnek ortalama değer

Bu verilen formülde birinci terim \bar{x} ile gösterilirse, \bar{x}' ü aşağıdaki gibi yazılmış olacaktır.

$$\bar{x}' = \bar{x} + \frac{n_2}{n}(\bar{x}_2' - \bar{x}_2) \quad [36]$$

\bar{x} : İlk seçilen örnekte yanıtlanmama olması durumunda elde edilecek örnek ortalaması

\bar{x}' 'nün varyansı aşağıdaki gibi tanımlanabilir.

$$V(\bar{x}') = E(\bar{x}' - \bar{X})^2$$

\bar{x}' yerine [36] ile verilen formül kullanılacak olursa, varyans formülü aşağıdaki gibi olur.

$$\begin{aligned} V(\bar{x}') &= E\left[\bar{x} + \frac{n_2}{n}(\bar{x}_2' - \bar{x}_2) - \bar{X}\right]^2 \\ &= E\left[(\bar{x} - \bar{X}) + \frac{n_2}{n}(\bar{x}_2' - \bar{x}_2)\right]^2 \end{aligned}$$

$$V(\bar{x}') = E(\bar{x} - \bar{X})^2 + 2E[(\bar{x} - \bar{X})\frac{n_2}{n}(\bar{x}_2' - \bar{x}_2)] + E\left[\frac{n_2}{n}(\bar{x}_2' - \bar{x}_2)\right]^2 \quad [37]$$

[37] eşitliğinin birinci terimi aşağıdaki gibi gösterilirken,

$$E(\bar{x} - \bar{X})^2 = \frac{N-n}{N.n} S^2 \quad [38]$$

eşitliğin ikinci terimi sıfırdır. Eşitliğin üçüncü teriminde \bar{X} terimi bir kez eklenip bir kez çıkaracak olursa,

$$\left(\frac{n_2}{n}\right)^2 E [(\bar{x}_2' - \bar{X}) - (\bar{x}_2 - \bar{X})]^2$$

eşitliğine ulaşılmış olacaktır. Buradan üçüncü terim,

$$\begin{aligned} E\left[\frac{n_2}{n}(\bar{x}_2' - \bar{x}_2)\right]^2 &= \left(\frac{n_2}{n}\right)^2 [E(\bar{x}_2' - \bar{X})^2 - 2E(\bar{x}_2' - \bar{X})(\bar{x}_2 - \bar{X}) + E(\bar{x}_2 - \bar{X})^2] \\ &= \left(\frac{n_2}{n}\right)^2 [E(\bar{x}_2' - \bar{X})^2 - E(\bar{x}_2 - \bar{X})^2] \\ &= \left(\frac{n_2}{n}\right)^2 \left(\frac{1}{n_2'} - \frac{1}{n_2}\right) S_2^2 \\ &= \frac{1}{n} \left(\frac{n_2}{n}\right) \left(\frac{n_2}{n_2'} - 1\right) S_2^2 \end{aligned} \quad [39]$$

formülüne dönüşmüş olur. [39] formülünde,

$$\left(\frac{n_2}{n}\right) \left(\frac{n_2}{n_2'} - 1\right) = k'$$

ile gösterilecek olursa, bu formül k' cinsinden aşağıdaki gibi tanımlanabilecektir.

$$= \frac{1}{n} k' S_2^2$$

[40]

Burada, $k' > 0$ 'dır. $k' = 0$, $n_2' = n_2$ ve altörnekleme yapılmamış demektir.

[37]varyans formülü, [38], [40] ve ikinci terim sıfır olduğu sonuçlarına dayalı olarak, aşağıdaki şekilde tanımlanmış olacaktır.

$$V(\bar{x}') = \left(\frac{N - n}{Nn}\right) S^2 + \frac{k' S_2^2}{n}$$

3.3.2. n , n_2' ve k' 'nin optimum değerleri

Srinath'ın yöntemiyle, yanıtlamama olması durumunda yığından alınması gereken optimum ilk örnek genişliği, yanıtlamayanlar arasından seçilecek altörnek çapı ve optimum altörnekleme kesri hesaplanmaya çalışılmıştır.

Bu bölümde ilk örnek genişliği n ve altörnek genişliğinin hesaplanması için gerekli k' 'nin hesaplanması için kurallar önerilmiştir. Bir önceki kesimde, k' aşağıdaki gibi tanımlanmıştı.

$$k' = \left(\frac{n_2}{n}\right) \left(\frac{n_2'}{n_2} - 1\right)$$

Buradan n_2' genişliği,

$$n_2' = \frac{n_2^2}{k'n + n_2} \quad [37]$$

olarak gösterir ve $k' > 0$ dir. $k' = 0$ olması, altörnekleme yapılmamış ve $n_2' = n_2$ demektir.

Altörnekleme kuralı altında, beklenen maliyet aşağıdaki şekilde tanımlanabilir.

$$E(C) = C_0n + C_1nW_1 + C_2E\left(\frac{n_2^2}{k'n + n_2}\right) \quad [42]$$

n ve k' 'nin optimum değerleri, beklenen maliyeti minimum yapan $V(\bar{x}') = \epsilon^2$ 'den hesaplanır. Buradan, [42] formülüyle verilen beklenen maliyet fonksiyonunun üçüncü terimi n ortak parantezine alınırsa, formül;

$$E(C) = C_0n + C_1nW_1 + C_2E n \left(\frac{\frac{n_2^2}{n}}{k'n + \frac{n_2}{n}} \right)$$

şekline dönüşür. Burada,

$$E\left(\frac{n_2}{n}\right) = E(w_2) = W_2$$

olduğundan, maliyet fonksiyonunun beklenen değeri,

$$E_1(C) = C_0n + C_1nW_1 + C_2E \cdot \frac{nW_2^2}{k' + W_2} \quad [43]$$

şeklinde yazılabilir. $E_1(C)$ maliyet fonksiyonunun k'' 'ye göre türevi alınıp sıfıra eşitlendiğinde k'' 'nin optimum değeri aşağıdaki gibi bulunmuş olacaktır.

$$k'_{opt} = \sqrt{\frac{(S^2 - W_2S_2^2)c_2W_2^2}{S_2^2(c_0 + c_1W_1)}} - W_2 \quad [44]$$

k'_{opt} formülü, Cochran'ın yönteminde elde edilmiş olan [28] denklemindeki k_{opt} 'a göre tanımlanacak olursa, aşağıdaki gibi olur.

$$k'_{opt} = (k_{opt} - 1) W_2$$

n ; yanıtlamama olması durumunda yığından alınacak örnek çapı, aşağıda verilmiş olan formülden hesaplanır.

$$n = \hat{n} \left(1 + \frac{k'S_2^2}{S^2} \right) \quad [45]$$

\hat{n} : Tam yanıt olması durumunda yığından alınması gereken örnek genişliği,

3.3.3. $L > 1$ için altörnekleme kuralı

El-Badry (43) tarafından önerilen örnekleme işlemi süreci şöyledir:

1. Yığımdan tesadüfi seçilen n çaplı örnek birimlerine anket formları postalanır,
2. Posta anketi geri dönüşüm süresinin sonunda n_1 yanıtlayanlar ve n_{12} yanıtlanmayanlar olduğu gözlenmişse, n_{12} yanıtlanmayanlar örneğinden $n'_{12} = n_{12} / k_1$ çaplı bir altörnek tesadüfi seçilerek, n'_{12} birimlerine tekrar anket formları postalanır,
3. Gönderilen anket formlarını n_2 kişinin yanıtladığı, n_{22} kişinin de yanıtlamadığı varsayıldığında, tekrar n_{22} yanıtlanmayanlar grubundan tesadüfi seçilen $n_{22}' = n_{22} / k_2$ çaplı altörnek birimlerine üçüncü bir girişim olarak anketler postalanır ve böylece süreç devam ettirilir,
4. Böylece, $(L + 1)$ inci girişim, n_{L2} yanıtlanmayanlarından seçilen $n'_{L2} = n_{L2} / k_L$ çaplı altörneğe uygulanır,
5. $(L + 1)$ inci girişimde hiç yanıtlanmayan birim kalmamıştır.

Srinath, \bar{X} yığın ortalamasının yansız bir tahminini aşağıdaki gibi tanımlamıştır.

$$\bar{x} = \frac{1}{n} (x_1 + k_1 x_2 + k_1 k_2 x_3 + \dots + k_1 k_2 k_3 \dots k_{L+1} x_{L+1}) \quad [46]$$

Burada, x_1, x_2, \dots, x_{L+1} sırasıyla n_1, n_2, \dots, n_{L+1} yanıtlayıcılarına ait değişken değerleri, $k_1, k_2, k_3, \dots, k_{L+1}$ ise 1'den büyük sabitlerdir.

\bar{x} 'nin varyansı aşağıdaki şekilde tanımlanmıştır.

$$V(\bar{x}) = \left(\frac{N-n}{Nn}\right)S^2 + \frac{(k_1-1)W_{12}S_{12}^2}{n} + \frac{1}{n} \sum_{j=2}^L \left(\prod_{i=1}^{j-1} k_i\right)(k_j-1)W_{j2}S_{j2}^2 \quad [47]$$

$V(\bar{x}') = \epsilon^2$ olması için gerekli minimum n , belirlenen $k_1, k_2, k_3, \dots, k_L$ değerleri ile aşağıdaki gibidir.

$$n = \hat{n} \left(1 + \frac{(k_1-1)W_{12}S_{12}^2}{S^2} + \frac{1}{S^2} \sum_{j=2}^L \left(\prod_{i=1}^{j-1} k_i\right)(k_j-1)W_{j2}S_{j2}^2 \right) \quad [48]$$

Burada, $k_1, k_2, k_3, \dots, k_L$ değerlerinin sabit tutulmasıyla $W_{12}, W_{12}, \dots, W_{L2}$ nin farklı değerleri için farklı n değerleri karşılık gelecektir.

Yanıtlamama olması durumu için, örnekleme kurallarına alternatif bir yöntem olarak geliştirilen bu altörnekleme yönteminde aşağıdaki tanımlamalar yapılmıştır.

$$n_{12}' = \frac{n_{12}^2}{k_1'n + n_{12}} \quad [49]$$

$$n_{j2}' = \frac{\prod_{i=1}^j n_{i2}^2}{k_j' n \prod_{i=1}^{j-1} n_{i2}'^2 + n_{j2} \prod_{i=1}^{j-1} n_{i2}^2}, \quad j = 2, 3, \dots, L \quad [50]$$

k_1', k_2', \dots, k_L' 'nin sıfırdan büyük sabitler olduğu kabul edilmiştir. [46] formülü,

[49] altörnekleme kuralı altında \bar{X} 'nin yansız bir tahmini olarak ele alınacaktır. Bu

altörnekleme kuralı altında, \bar{x} 'nin varyansı aşağıdaki gibi gösterilebilir.

$$V(\bar{x}) = \frac{N-n}{Nn} S^2 + \frac{1}{n} \sum_{j=1}^L k_j' S_{j2}^2 \quad [51]$$

[51] formülü $W_{12}, W_{12}, \dots, W_{L2}$ değerlerinden bağımsızdır. $V(\bar{x}') = \epsilon^2$ için gerekli, sabit $k_1', k_2', k_3', \dots, k_L'$ değerleri ile, n ilk örnek genişliği aşağıdaki gibi verilir.

$$n = \hat{n} \left(1 + \frac{1}{S^2} \sum_{j=1}^L k_j' S_{j2}^2 \right) \quad [52]$$

[51] formülünde olduğu gibi, bu formülde de W_{j2} , $j=1,2,\dots,L$ değerleri yoktur. Optimum k_j' , $j=1,2,\dots,L$ ve n değerleri, verilen bir tahmin edicinin varyansı ile beklenen maliyeti minimum yapılacak şekilde de elde edilebilir.

Bu durumda, W_{j2} , $j=1,2,\dots,L$ değerleri kümesi varsayılr. k_j' , $j=1,2,\dots,L$ ve n , tahmin edicinin istenilen doğru değerini elde edebilmek için hesaplanır. W_{j2} 'nin varsayılan değeri ile gerçek değeri eşit olduğunda yapılan tüm işlemler aynı optimum maliyet ve istenilen doğruluğu verir. Eğer eşitlik sağlanamamışsa, tahmin edicinin, [49] denkleminde altında, en azından istenilen doğruluğuna ulaşılabilir ve yığındaki gerçek yanıtlamama oranları arasındaki farklılığa dayanan optimumluk, daha yüksek bir maliyetle gerçekleştirilebilir. Sinath (11), $L>1$ için altörnekleme kuralını bir örnekle vermeye çalışmıştır.

Örnek. Birimlerden veri toplamak amacıyla, ilk iki girişimin posta ve üçüncü girişimin yüz-yüze görüşme anketi yöntemiyle yürütüldüğü, üç girişimin yapıldığı varsayımı yapılmıştır. Yığından seçilen örnek çapı $n=1000$ ve $S^2 = 0,45$, $S_{12}^2 = 0,16$, $S_{22}^2 = 0,09$ olarak kabul edilmiştir. Ayrıca maliyetler, Hansen ve Hurwitz'te olduğu gibi, $C_0= 0,10\$$, $C_1=0,40\$$, $C_2=4,50\$$ olarak alınmıştır. Burada, C_1 : Birinci ve ikinci tabakadan birim başına veri elde etme ve işleme maliyetini, C_2 : Üçüncü tabakada birim başına veri elde etme ve işleme maliyetini ifade etmektedir.

Aşağıda verilen Çizelge 3.1. çizelgesi, çeşitli yanıtlamama oranları için optimum örnek çapı, üç farklı beklenen maliyet durumlarını göstermektedir.

Çizelge 3. 1. Beklenen maliyetlerin kıyaslanması

Gerçek ağırlıklar		Optimum örnek çapı	Optimum beklenen maliyet	$W_{12}=0,4$ ve $W_{22}=0,3$ olarak varsayıldığında, (3.3.15) kuralı altında beklenen maliyet (\$)	Yanıtlamayanların %100 örneklendiği durum için beklenen maliyet (\$)
W_{12}	W_{22}				
(1)	(2)	(3)	(4)	(5)	(6)
,80	,70	2304	1334	1572	3450
,80	,50	2406	1051	1149	2630
,80	,30	1785	800	843	1810
,70	,60	2024	1198	1336	3030
,60	,50	1801	1067	1131	2610
,60	,20	1457	701	730	1380
,50	,40	1606	941	957	2190
,40	,30	1440	820	820	1770
,30	,20	1289	707	726	1350
,20	,10	1153	601	676	930

$W_{12}=0,40$ ve $W_{22}=0,30$ olarak varsayıldığında, önceden belirlenen bir varyansla optimum beklenen maliyette, k_1' ; k_2' ve n 'in optimum değerleri sırasıyla 1,57; 4,81; 1440 dir. Çizelge, yanıtlamamanın çeşitli oranları için, optimum beklenen maliyetleri ve posta anketlerinin gönderildiği 1440 kişiden [49] kuralı altında $k_1'=1,57$ ve $k_2'=4,81$ ile altörneklerin seçildiği beklenen maliyetleri göstermektedir. Ayrıca çizelgenin son sütununda, $n=1000$ kişiye posta anketi uygulama ve posta anketini yanıtlamayanların tümünden yanıt alınmak üzere takiplerin yapılması durumu için beklenen maliyetleri gösterilmiştir. 5. sütuna bakıldığında, maliyet değerlerinin 4'üncü sütundaki değerlerden çok da yüksek olmadığı görülmektedir.

3.4. Deming'in Yöntemi

Deming (12)'de, çift örnekleme yönteminin, maliyet ve hata modelleme yaklaşımlarına uygun düştüğünü savunmuştur. Örnek birimleriyle her bir görüşme yapılarak veri toplanılması aşamasında maliyetlerin farklılık gösterdiği kabul edilmiştir. Ayrıca, sonuç istatistiklerinin örnekleme ve yanıtlamama hatası, ikinci aşama örnekleme düzenine ve tüm yanıtlamayanlar örneğinin ölçümündeki başarıya dayandığı belirtilmiştir.

Yanıtlamayanların varlığında, yeniden görüşme sağlanması, bunun için ise; hata kareler ortalamasını minimum yapan bir çift örnekleme planı tanımlanmıştır. Politz ve Simmons (45), Birnbaum ve Sirken (46), Hansen ve Hurwitz (9) ve başka stratejiler, araştırmanın başlangıcındaki maliyetler ve hatalar bilgisini içeren Deming planının özelliklerini paylaşmaktadır. Ancak, çoğu, yanıtlamamanın iki farklı kategorisindeki farklı maliyet ve hata özelliklerine sahip redler ve ulaşılamayanlar arasındaki farklılıkları ihmal etmiştir (1).

Deming (12)'in yaklaşımı;

1. Anketörün yapmış olduğu her bir yeniden görüşme için farklı yanıt olasılıklarına sahip gruplar olduğu,
2. Yapılan ilk yeniden görüşme sonrasında yanıtlamayanlardan altörnek alınması,
3. Yanıtlamayanlar arasından tesadüfi seçilen tüm altörnek birimlerden yanıt alınması (yanıt alınıncaya kadar yapılacak yeniden görüşmeleri içerir)

varsayımlarını yapmıştır (44).

Deming'in yöntemiyle de, yanıtlamayanlar örneğinden alınacak altörnek çapı ve yığına ilişkin ortalamanın tahmini belirlenmeye çalışılmıştır. Bu, yanıtlayanlarla yapılan ilk yeniden girişimlerde yanıt alınamayan birimlerden ne kadarıyla yeniden görüşme yapılacağına ve ne kadarının görüşme dışı tutulacağına karar verilmesi sürecine dayalı olarak yürütülmeye çalışılmıştır.

Karar verme aşaması, yeniden görüşme modelleri bilgisini, araştırma değişkeninde gruplar arasındaki (her bir yeniden görüşmelerde yanıtlayanlar grupları) ortalamalardaki farklılıklar bilgisini ve gruplar arasındaki birim varyanslardaki farklılıklar bilgini gerektirmektedir. Burada, araştırmanın tümünü yapmak için gereken toplam maliyet sabit kabul edilmiştir.

Her bir örnek birimiyle yapılan her yeniden görüşme için, 6 yanıt olasılığından (0; 0,125; 0,250; 0,500; 0,750 veya 1,000 olasılıklarından) birine sahip olduğu varsayımı yapılmıştır. Bu olasılıklar, yapılan yeniden görüşmelerde, her birim için sabittir (1).

Örneğin, bazı kişilerle vardır ki, o kişilerle görüşme yapılması olasılığı her zaman için 0,250 dir. Bazı kişilerde vardır ki, bu kişiler kendileriyle her zaman görüşme yapılıns isterler. Bu kişiler için yanıt olasılığı 1,000'dir. Ayrıca, kesinlikle görüşme yapılmasını istemeyen, görüşmeleri sürekli reddeden kişiler olduğu belirtilmiştir (yanıt olasılığı 0 veya sıfıra yakın bir değer).

Bu yöntemde, yeniden görüşmede örnek ortalaması; yapılan bir yeniden görüşme ile bir önceki görüşme girişiminden elde edilen ortalamaların ağırlıklandırılmış bir ortalaması olarak ele alınmıştır.

Dolayısıyla, yapılan üç yeniden görüşme sonundaki örnek ortalamasının tahmini aşağıdaki gibi olacaktır.

$$\bar{y}_{1+2+3} = w_1 \bar{y}_1 + w_2 \bar{y}_2 + w_3 \bar{y}_3 \quad [53]$$

w_1 : *h. yeniden görüşmede görüşme yapma olasılığı*

\bar{y}_h : *h. girişimde yanıt verenlerden hesaplanan ortalama değeri, h=1,2,3*

Bu model, her bir girişimde elde edilen durumların oranlarıyla ağırlıklandırılmış her bir girişimde yapılan görüşmelerin ortalamalarının basit doğrusal bir birleşimidir. Girişim sayısı arttıkça modele eklenecek ortalamaların sayısı da artacaktır.

Deming (12), ampirik bir veri ile Çizelge 3.2 ve Çizelge 3.3'te çift örnekleme yöntemini bir örnekte incelemeye çalışmıştır. Burada girişim maliyetleri ilk girişim için 3,00\$, her bir alt girişim için ise 5,00\$ olarak belirlenmiştir. Bu durumda yapılacak ilk girişim için toplam maliyet 3,00\$, iki kez yapılacak girişim için toplam

maliyet 8,00\$ olur. Bu maliyetler görüşme sağlanmış veya görüşme sağlanamamış (başarısız) durumların her ikisi içinde geçerlidir.

Çizelge 3. 2. Deming modeli için göstermelik parametre değerleri

Yanıt Olasılıkları Sınıfı	Sınıflardaki Yığın Oranı	Araştırma Değişkeninin Ortalama Değeri	Araştırma Değişkeninin Birim Varyansı
0- (hiç görüşme yapılmamış)	0,05	2,25	
0,125	0,10	2,00	2,00
0,250	0,10	1,75	1,75
0,500	0,20	1,50	1,50
0,750	0,25	1,25	1,25
1,00	0,30	1,00	1,00
Toplam	1,00	1,40	

İlk girişimden sonra örnekteki yanıtlamayanlar optimal kesri, yanıtlamayanlar altörneğinde bulunan birimlerle yeni bir görüşme yapılması, iki görüşme yapılması, üç görüşme daha yapılması veya ne yapılıp yapılmayacağı kararı üzerine koşulludur. Optimal çift örnekleme kesrinin hesaplanabilmesi için bir maliyet ve hata modeline ihtiyaç vardır.

Araştırmanın toplam maliyeti aşağıdaki gibi tanımlanmıştır.

$$C = C_1 n + C_2 f (n-r) \quad [54]$$

- C_1 : İlk aşama örneğinde birim başına düşen maliyet, yalnızca ilk görüşme maliyetleri
 n : Toplam örnek çapı
 C_2 : İkinci aşamada yanıtlamayanlar için çeşitli sayıda seçilen girişimlerin görüşme başına maliyeti
 f : Yapılan bir girişimden sonra takibi yapılacak yanıtlamayanlar kesri
 r : İlk aşamada yanıtlayanların sayısı

f optimal kesri, maliyet oranlarına, varyansla ve yan terimiyle orantılı olarak bulunur, bu nedenle de hata kareler ortalaması şöyle yazılabilir:

$$HKO(\bar{y}) = A + \frac{B}{n} + \frac{C}{fn}$$

A : Yan terimi (hiç bir zaman yanıtlamayacak birimlerle, ve diğer ikinci aşamada seçilen ancak tüm girişimlere rağmen ulaşılamayan birimleri belirtir.)

B,C : Sırasıyla birinci ve ikinci aşama görüşmeleri için birim varyans fonksiyonları

Buradan, optimal $f; \frac{CC_1}{BC_2}$ teriminin kareköküne eşittir. İlk girişimde hata kareler ortalaması için A aşağıdaki gibidir.

$$A = E(\bar{y}_1) - \bar{y}_1$$

Burada, \bar{y}_1 ; ilk yeniden görüşmelerindeki ortalama, \bar{y}_1 ise; 0,00 yanıtlanma olasılığına sahip örnek durumları dışındaki tüm durumlara dayalı örnek ortalamasıdır.

Eğer grup ortalamaları, varyanslar ve katılımcı olabirlikleri Çizelge 3.2'deki gibi, $C_1=3,00$ \$, $C_2=5,00$ \$ ve maksimum yedi girişim yapılması kısıtı getirilmiş ise, optimal ikinci aşama kesri; yaklaşık 0,6 olarak bulunur. Bu, ilk girişimden sonra, örnek birimlerinin yüzde 60'ının yanıtlamayanlardan oluştuğunu, çoğuna da 6'dan fazla girişimin yapıldığı anlamına gelmektedir.

Yukarıda verilen örneğe dayalı olarak, eğer ilk olarak 1000 kişi örneğe seçilmiş ise ilk girişimde 625 görüşme yapılmış demektir. bu durumda, yanıtlamayan 375 kişiden 225 ($375 \times 0,6 = 225$) kişi ikinci aşama örneğine alınır. İkinci aşamada yapılan yedi girişimin sonunda 162 kişiden yanıt alınmış, 63 kişiden yanıt alınamamıştır. 63 kişiden yüzde 60'ıyla hiç bir şekilde görüşme yapılma şansı olmayacağını beklenir.

Çizelge 3. 3. İlk girişimden sonra yanıtlanmayanların optimal 0,60 çift örnekleme varsayımıyla araştırma sonuçları

ÖRNEK DURUMLARININ SAYISI									
Girişim Sayısı	Aktif Olan	Girişimde Görüşme Yapılanlar	Yanıtlanmayan Olarak Kalanlar	Birikimli Görüşme Sayısı	Toplam Maliyet	Durum Başına Maliyet	Görelî Yan	Görelî Hata Kareler Ortalaması GHKO	GHKO/Toplam Maliyet X 10 ⁶
1	1000	625	375	625	3000\$	4,80	-0,11	0,117	39,0
2	225	76	149	701	4125	5,88	-0,08	0,085	20,6
3	149	36	113	737	4870	6,61	-0,06	0,069	14,2
4	113	21	92	758	5435	7,17	-0,05	0,059	10,9
5	92	13	79	771	5896	7,65	-0,04	0,053	9,0
6	79	9	70	780	6290	8,06	-0,03	0,049	7,8
7	70	7	63	787	6638	8,43	-0,03	0,046	6,9

Çizelge 3.3, maliyet kısıtı altında, hataların oranının (GHKO/toplam maliyet) sürekli bir düşüşünü gösterir.

4.UYGULAMA

Nevşehir Tekel Fabrikası çalışanlarına, iş ortamı profilinin tespit edilmesi amacıyla bir anket çalışması uygulanmıştır. Anket formu birinci soru grubu ve ikinci soru grubu olmak üzere toplam 34 soru ve 6 demografik sorudan oluşmaktadır. Anket formunun birinci soru grubu, işletme çalışanları üzerinde örgütsel adaletin uygulanıp uygulanmadığını ölçen 21 soruyu içermektedir. İkinci soru grubu ise işletme çalışanlarının kendilerini örgüte bağlı hissedip hissetmediklerini ölçen 17 sorudan oluşmaktadır.

Uygulama çalışması birinci soru grubu ve ikinci soru grubu üzerinden yapılmıştır. Birinci ve ikinci soru grubunda son iki soru aynıdır.

Anket formu tasarımı yapıldıktan sonra, Tekel Fabrikası çalışanlarından veri toplamak üzere posta anketi yönteminin uygulanacağına karar verilmiştir. Posta anketi yönteminin tercih edilmesinin nedeni, posta anketlerine olan yanıt düzeyinin ne olacağının tespit edilerek, yanıtlamama olması durumunda çift örnekleme yönteminin bir uygulamasının yapılabilmesidir.

Posta anketi diğer yöntemlere kıyasla en düşük maliyetli ve en düşük yanıtlanma oranına sahip anket yöntemidir. Tekel Fabrikasında yürütülecek posta anketi çalışmasında yanıtlamama durumu söz konusu olması halinde, Hansen ve Hurwitz (9), Cochran (10), Srinath (11) ve Deming (12)'in yanıtlamama için çift örnekleme yönteminin bir uygulaması yapılabilmiş olacaktır.

Araştırma kapsamında Hansen ve Hurwitz'in, Cochran'ın ve Srinath'ın çift örnekleme yöntemleri uygulanmıştır. Uygulama çalışmasında Deming'in çift örnekleme yöntemi kullanılmamıştır, çünkü bu yöntemde göre kişilere ilişkin yanıt oranları bilgisine ihtiyaç vardır. Ancak bu çalışmada kişilere ilişkin yanıt oranlarının tahmini yapılamadığı için, Deming'in yönteminin de uygulaması gerçekleştirilememiştir.

Uygulama çalışmasında, ikiden fazla aşama altörnek alma işlemine gerek kalmamıştır, çünkü yanıtlamayanlar örneğinden seçilen altörnek birimlerinin tümünden yanıt alınmıştır. Eğer yanıtlamayanlardan tesadüfi olarak seçilen altörnek birimlerinin tümünden yanıt alınamamış olsaydı, üçüncü veya daha fazla aşama yeniden altörnek seçim işlemi yapılması gerekebilirdi.

Aşağıdaki bölümlerde sırasıyla n örnek çapının belirlenmesi, ikinci olarak n örnek birimlerine postalanan anket formlarının yanıtlanma durumları ve yanıtlanan posta anketlerine dayalı (tek örnekleme yöntemiyle) yapılan yığın parametre tahminleri, üçüncü olarak yanıtlamama olması durumunda yığın parametrelerinin çift örnekleme yöntemiyle tahmin edilmesi ve son olarak tek örnekleme ve yanıtlamama olması durumunda çift örnekleme yöntemi ile hesaplanan yığın parametresi tahminlerinin kıyaslaması yapılmıştır.

Uygulama çalışması, anket formunda yer alan birinci soru grubu ve ikinci soru grubu için ayrı ayrı yürütülmüştür.

Birinci ve ikinci soru grubunda yer alan sorular likert ölçeğinde hazırlanmıştır. Her bir soru 5'li likert ölçeğinde hazırlanmış olup, sorular her bir soruya kişilerin aşağıda yer alan yanıt seçeneklerinden kendilerine en uygun olan ifadeyi seçmeleri istenmiştir. Sorulara ilişkin verilen yanıtlar aşağıdaki gibi tanımlanarak kodlanmış ve veri girişi SPSS paket programında buna göre yapılmıştır.

1: Kesinlikle aynı fikirde değilim

2: Aynı fikirde değilim

3: Kararsızım

4: Aynı fikirdeyim

5: Kesinlikle aynı fikirdeyim

4.1. Birinci Soru Grubu

Tekel Fabrikası personeli sayısı 384'tür. Örnekleme yapılmak üzere yığından basit tesadüfi örnekleme yöntemiyle n çaplı örnek seçimi yapılarak, bu n birime hazırlanmış olan anket formları postalanmıştır. Aşağıdaki bölümlerde örnekleme süreci verilmeye ve yığın parametrelerine ilişkin istenen tahminler yapılmaya çalışılmıştır.

4.1.1. n örnek çapının belirlenmesi

Araştırmaları yapmak üzere yığından yeter büyüklükte örnek seçmek ve seçilen örnek genişliğinin en uygun büyüklük olduğunu söyleyebilmek gerekir. Uygunluk, iki ölçü ile ölçülebilir bunlardan biri tahminin standart hatası üzerine konulacak sınırlar, diğeri ise maliyettir (37).

Burada en uygun örnek büyüklüğü, standart hata üzerine konulan sınırdan hesaplanmaya çalışılmıştır. Standart hata üzerine konulan sınırlar d olup, istenilen en uygun örnek çapı aşağıda verilen [54] formülden hesaplanır.

duyarlılık = (güvenilirlik) x (standart hata)

$$d = z \cdot sh(\bar{x}) \quad [54]$$

d : Duyarlılık
z : Güvenilirlik katsayısı
sh(\bar{x}) : Ortalamanın standart hatası

Örnek ortalamasının belirlenen bir güvenilirlik düzeyinde yığın ortalamasının $\pm d$ sınırları içinde olması, örnek tahmininin duyarlılığının belirlenen güvenilirlik düzeyinde d sınırları $\bar{x} \pm z \cdot sh(\bar{x})$ sınırları içinde olması demektir (20).

sh(\bar{x}) aşağıdaki gibidir.

$$sh(\bar{x}) = \sqrt{\frac{N-n}{N} \frac{S^2}{n}}$$

Buradan [54] eşitliği,

$$d = z \cdot \sqrt{\frac{N-n}{N} \frac{S^2}{n}}$$

olarak yazılabilir.

Örnek çapının tahmin edilebilmesi için, S^2 yığında birim başına düşen varyansının biliniyor olması gerekir. S^2 bilinmediğinden yığından n_0 çapında bir ön örnek seçilerek, bu ön örnekten S^2 'nin tahmini yapılarak istenen asıl n örnek çapı hesaplamasına geçilebilir.

Eğer n örnek çapı $> n_0$ ise, aradaki fark tamamlanmalı ve yığından $n-n_0$ genişliğinde örnek seçimi daha yapılmalıdır.

n 'i belirleyebilmek ve anket formlarında yanıtlanmayan veya anlaşılmayan soru veya konuların tespit edebilmek amacıyla yığından tesadüfi olarak $n_0=40$ çaplı ön örnek alınarak bir öntest çalışması yürütülmüştür. Listede bulunan kişilere 1'den 384'e kadar numara verilerek, bu numaralar içinden eşit seçilme olasılığı ile ve yerine koymadan seçim yöntemiyle 40 numara seçilmiştir, seçilen bu numaralara karşılık gelen kişiler de örneğe alınmıştır.

40 kişi üzerinden yapılan öntest çalışması sonucunda aşağıdaki istatistiksel değerlere ulaşılmıştır.

Çizelge 4. 1. Öntest sonucunda birinci soru grubundan elde edilen istatistik değerleri

Ortalama (\bar{x})	=	3,6245
Ortalamanın standart hatası ($sh(\bar{x})$)	=	0,194
Mod	=	4
Standart Sapma (s)	=	1,2965
Varyans (s^2)	=	1,6809

Öntest çalışmasında, $40 \times 21 = 840$ veri üzerinden yığın varyansı tahmin edilmiş ve $s^2 = 1,6809$ olarak bulunmuştur.

$d = z \cdot sh(\bar{x})$ formülünde %95 güvenilirlik düzeyinde $z=2$ ve $d=0,09$ olarak alınmıştır. Buradan yığından seçilecek n örnek büyüklüğü hesaplanabilir.

$$d = z \cdot sh(\bar{x})$$

$$0,09 = 2 \cdot \sqrt{\frac{N-n}{N} \frac{s^2}{n}}$$

$$(0,09)^2 = 4 \cdot \frac{384-n}{384} \frac{1,68}{n}$$

$$n = 262$$

Anket çalışmasını yürütmek üzere $N=384$ çaplı yığından seçilecek örnek büyüklüğü 262 olarak tespit edilmiştir.

4.1.2. Tek örnekleme yöntemi

Posta anketlerinin $N=384$ çaplı yığından seçilecek 262 kişiye postalanacağına karar verilmiştir. Öntest çalışmasıyla yığından tesadüfi seçilen 40 kişiyle görüşme yapılmıştı dolayısıyla anket çalışması yığından tesadüfi seçilecek $262-40=222$ kişiye daha uygulanacaktır ($n_0=40$ olduğu için $262-40$ genişliğinde örnek daha seçilmiştir). Listede bulunan kişilere 1'den 384'e kadar numara verilerek, bu numaralar içinden eşit seçilme olasılığıyla ve yerine koymadan seçim yöntemiyle 222 numara seçilmiştir, seçilen bu numaralara karşılık gelen kişiler de örneğe alınmıştır. Tekel Fabrikasında çalışan 384 kişinin adres ve telefon bilgilerinin bulunduğu listeden tesadüfi seçilen 222 kişinin adreslerine anket formları postalanmıştır.

Anket formları kişilerin adreslerine postalanmış ve geri dönüşüm süresi olarak belirlenen 2 hafta sonunda yanıtlamayanlara anket çalışmasını yanıtlamaları için

hatırlatmalar yapılmıştır. Ancak yapılan hatırlatmalar sonucunda anketi yalnızca 30 kişi yanıtlamıştır. Yanıtlayan kişi sayısı çok düşük olmuştur.

Öntest ile 40 kişiyle görüşme yapılmıştı dolayısıyla yığından tesadüfi seçilen birimlerden anket çalışmasına katılanların sayısı yalnızca $40+30 = 70$ olmuştur. Yanıtlamayanların sayısı $262-70=192$ kişidir. Çizelge 4.2., 70 kişiye ilişkin istatistik değerlerini göstermektedir.

Çizelge 4. 2. Yanıtlayanlar tabakasına dayalı hesaplanan istatistik değerleri

<i>Ortalama (\bar{x})</i>	=	3,6960
<i>Ortalamanın standart hatası ($sh(\bar{x})$)</i>	=	0,1293
<i>Mod</i>	=	4,00
<i>Birim Başına Düşen Standart Sapma (s)</i>	=	1,1965
<i>Birim Başına Düşen Varyans (s^2)</i>	=	1,4315

Tek örnekleme yöntemiyle, yığın ortalamasının ve örnek ortalaması varyansının tahmini, posta anketini yanıtlayan 70 kişiden elde edilen verilere dayalı olarak yapılmıştır.

4.1.2.1. Yığın ortalamasının tahmini

Posta anketini yanıtlayan 70 çalışana ilişkin istatistik değerleri yukarıdaki çizelgede verilmişti. 70 kişiye yöneltilen birinci soru grubunda 21 soru ve 70 kişiye ilişkin toplam $70 \times 21 = 1470$ veri vardır. 1470 veri üzerinden yığın ortalamasının tahmini $\bar{x} = 3,6960$ olarak hesaplanmıştır.

Yığın ortalamasının tek örnekleme yöntemiyle tahmininden ($\bar{x} = 3,6960$), %95 güvenilirlik düzeyinde kişilerin genelinin çalıştıkları kuruma bağlılıklarının yüksek olduğu yorumu yapılabilir.

4.1.2.2. \bar{x} 'nin varyans tahmini

Yığın varyansının tahmini, Çizelge 4.2.'de gösterildiği üzere, $s^2=1,4315$ olarak bulunmuştur. Örnek ortalamasının varyansı yalnızca yanıtlayanlara bağlı olarak hesaplanacak olursa, $n=70$ ve varyans,

$$V(\bar{x}) = \frac{N-n}{N} \cdot \frac{S^2}{n}$$

formülünden bulunur. Burada S^2 yığın varyansının tahmini $s^2=1,4104$ olarak alındığında, örnek ortalamasının varyans tahmini aşağıdaki gibi bulunur.

$$v(\bar{x}) = \frac{384-70}{384} \cdot \frac{1,4315}{70}$$

$$v(\bar{x}) = 0,0167$$

$n=70$ iken, \bar{x} 'ya ilişkin varyans değeri $v(\bar{x}) = 0,0167$ oldu. Örnek ortalamasının varyans tahmini küçük bulunmuştur. Ancak acaba varyansın düşük bulunması yapılan tahmine olan ilgili duyarlılığa erişmek bakımından bizi tatmin etmekte midir?

Yürütülen örnek çalışmasında yanıt oranı %27 olmuştur. Düşük yanıt oranı olması yığına ilişkin yapılan tahminlere olan güveni azaltmaktadır. Tüm araştırmalarda örneğe seçilen tüm birimlerin anket çalışmasını yanıtlaması istenir. Posta anketleri düşük maliyetli olması nedeniyle tercih edilirken, yürütülen posta anketlerinin çoğunluğunda yanıt oranları düşük bulunmaktadır.

Acaba, %27'lik düşük yanıt oranı olması durumu için yığından seçilmeyen 384-262=192 kişiden tesadüfi seçim yapılarak, bu kişilere posta anketleri mi postalanmalı, yoksa yanıtlamayan birimlerinden bilgi almak amacıyla daha farklı veri toplama tekniği veya farklı bir yöntem mi kullanılmalıdır?

Yanıtlamama olması durumunda, yanıtlayanların yanıtlamayanlardan farklı olduğu varsayımından, yanıtlamayan birimlerinden veri elde edilmeye çalışılmalıdır. Yanıtlamayan birimlerinden veri toplama işlemi posta anketi yöntemiyle sağlanamadığı için, farklı bir veri toplama yöntemi kullanılması gerektiğine karar verilmiştir. Yanıtlamayanlardan veri toplama işleminin daha yüksek yanıt alınmasına imkan sağlayan yüz-yüze görüşme anketi yöntemiyle olacağına karar verilmiştir.

Yanıtlamama olması istenilmeyen bir durumdur, düşük yanıt oranı yapılan tahminlere olan güveni azaltacağından, yanıtlamamanın giderilmesi veya en aza indirgenmesi için önerilen yöntemlerden biri olan çift örnekleme yöntemi kullanılmıştır.

4.1.3. Yanıtlamama için çift örnekleme yöntemi

Cevaplandırılmak üzere anket formları gönderilen kişilerin %73 gibi büyük bir çoğunluğu anket formlarını yanıtlamamıştır. Yanıtlamayanlar tabakasından yanıt alınmak üzere Tekel İçki Fabrikasına gidilmiş, yanıtlamayanların ağırlıkla Tekel Fabrikası işçileri olduğu, yanıtlayanların ise ağırlıkla Tekel Fab. Memurları ve bölüm sorumluları olduğu tespit edilmiştir. Buradan, yanıtlayanların yanıtlamayanlardan farklı özelliklere sahip olduğu yorumu yapılabilir.

Örneğe seçilen birimlerden yanıtlamayanların tümüyle yapılacak yüz-yüze görüşme anketinin maliyetli, zaman alıcı ve daha çok emek harcanması demek olacağından, yanıtlamayanlardan altörnek alınması önerilmiştir.

$n=262$ kişi üzerinden yürütülen posta anketi çalışmasının yalnızca 70'inden yanıt alınabilmiş, 192 kişiden yanıt alınamamıştır.

Araştırma kapsamında Hansen ve Hurwitz'in, Cochran'ın ve Srianth'ın çift örnekleme yöntemleri uygulanmıştır. Yanıtlamayalar tabakasından tesadüfi olarak seçilecek altörnek çapını (n_2') ve yanıtlamama olması durumunda yığından seçilmesi

gereken ilk örnek çapının (n) optimum değerleri bulunarak, çift örnekleme yöntemiyle yığına ilişkin ortalama ve varyans tahmin edilmeye çalışılacaktır.

4.1.3.1. Hansen ve Hurwitz'in çift örnekleme yöntemi

Bu bölümde istenilen tüm hesaplamalar N ve N₂'nin yeteri kadar geniş olduğu ve

$\frac{N}{N-1} = 1$ ve $\frac{N_2}{N_2-1} = 1$ olduğu varsayımı ile $\sigma^2 = \sigma_2^2$ ele alınan yığının özellikleri

ile yanıtlamayanların oluşturduğu yığının özelliklerinin aynı olduğu varsayımı altında yürütülmüştür (σ^2 : Yığın varyansı, σ_2^2 : Yanıtlamayanlar tabakasının varyansı) .

4.1.3.1.1. k, n₂' ve n'in optimum değerlerinin hesaplanması

n örnek çapı yanıtlamama durumunu göz önüne almadan n= 262 olarak saptanmıştır.

Hansen ve Hurwitz, yanıtlamama olması durumunda yığından seçilmesi gereken optimum örnek çapını aşağıdaki formülden hesaplanmasını önermiştir.

$$n_{OPT} = \hat{n} [1 + (k-1)q]$$

Optimum altörnekleme kesri ise;

$$k_{OPT} = \sqrt{\frac{c_2 P}{c_0 + c_1 P}}$$

n_{OPT} : Yanıtlamama olması durumunda yığından alınacak optimum n örnek çapı

k_{opt} : Altörnekleme kesri

\hat{n} : Tam yanıt olması durumundaki örnek çapı

q : Yanıtlamayanlar tabakasının örnek oranı

formülünden hesaplanır.

Hansen ve Hurwitz'te (9), örnekleme yapma maliyeti,

$$C = c_0n + c_1n_1 + c_2n_2'$$

- C : Örnekleme yapma maliyeti
 c_0 : Birim başına anket formu postalama maliyeti
 c_1 : Yanıtlanan bir posta anketinin işlenilmesi maliyeti
 c_2 : Birim başına yüz-yüze görüşme yapılması ve alınan sonuçların işlenilmesi maliyeti
 n_2' : Yanıtlamayanlar tabakasından alınacak altörnek çapı

olarak verilmiştir.

Tekel fabrikasında yürütülen yüz-yüze görüşme anketi yöntemi için belirlenen maliyet değerleri aşağıdaki gibi verilmiştir.

$$\begin{aligned}
 c_0 &= 400\ 000\ TL \\
 c_1 &= 1500\ 000\ TL \\
 c_2 &= 5000\ 000\ TL
 \end{aligned}$$

Yığından tesadüfi olarak seçilen $n=262$ kişiye posta anketi gönderilmiş, anketleri yanıtlayan 70 kişi olmuştur.

Yanıtlayanlar	Yanıtlamayanlar
$n_1 = 70$	$n_2 = 192$
$p = \frac{n_1}{n} = \frac{70}{262} = 0,27 = w_1$	$q = \frac{n_2}{n} = \frac{192}{262} = 0,73 = w_2$

- n_1 : Yanıtlayanlar tabakasının örnek çapı
 n_2 : Yanıtlamayanlar tabakasının örnek çapı
 n : Örnek çapı
 p : Yanıtlayanlar tabakasının örnek oranı
 q : Yanıtlamayanlar tabakasının örnek oranı

Buradan, $p=0,27$, $q=0,73$ olarak belirlenmiştir. n_{OPT} bulabilmek için optimum altörnekleme kesri k 'nın hesaplanması gerekir. $\frac{N}{N-1}=1$, $\frac{N_2}{N_2-1}=1$ ve $\sigma^2=\sigma_2^2$ varsayımı altında optimum k 'da yukarıda verilen maliyet değerlerini optimum k 'da yerine yazarsak,

$$k_{OPT} = \sqrt{\frac{c_2 p}{c_0 + c_1 p}} = \sqrt{\frac{5000000 \times 0,27}{400000 + 1500000 \times 0,27}}$$

$$k_{opt} = 1,29$$

değerine ulaşılmış olur. Buradan, yanıtlamayanlar tabakasından alınacak optimum altörnek çapı aşağıdaki gibi bulunur.

$$n_2' = n_2 / k_{OPT}$$

$$n_2' = 192 / 1,29 = 148$$

Yanıtlamayanlar tabakasından yüz-yüze görüşme anketi uygulanmak üzere seçilecek altörnek çapı 148 olarak belirlenmiştir.

n yanıtlamama olması durumunda yığından seçilmesi gereken optimum örnek çapı,

$$n_{OPT} = \hat{n} [1 + (k-1)q]$$

$$n_{OPT} = 262 \{1 + (1,29 - 1)0,73\}$$

$$n_{opt} = 317$$

şeklinde bulunur.

Yığından alınması gereken optimum ilk örnek genişliği, yanıtlamama olması durumunda 317 olarak belirlenmiştir. Ancak, Hansen ve Hurwitz birimlerin tümüne (317-70=247 kişiye daha) posta anketleri göndermek yerine, yanıtlamayan birimlerinden alınacak altörnek birimleriyle yüz-yüze görüşme yapılmasının maliyette fayda ve ekonomi sağlayacağını belirtilmiştir. Dolayısıyla yalnızca yanıtlamayan birimlerinde seçilecek 148 kişiyle görüşme yapılması maliyette fayda sağlayacaktır.

4.1.3.1.2. Yığın ortalamasının tahmini

Yanıtlamama olması durumunda yığın ortalamasının yansız bir tahminini çift örnekleme yöntemiyle aşağıdaki formülden hesaplanabilir.

$$\bar{x}' = \frac{n_1}{n} \bar{x}_1 + \frac{n_2}{n} \bar{x}_2'$$

\bar{x}' : Yığın ortalamasının yansız bir tahmini

\bar{x}_1 : n_1 yanıtlayanlar tabakasının ortalama değeri

\bar{x}_2' : Altörneğin ortalama değeri

n_1 : Örnek birimlerinden yanıtlayanların sayısı

n_2 : Örnek birimlerinden yanıtlamayanların sayısı

n : Örnek çapı

\bar{x}' formülünde \bar{x}_2' bilinmeyendir. \bar{x}_2' 'ü altörnek ortalamasını bulabilmek için, yanıtlamayanlar arasından seçilecek $n_2'=148$ altörnek birimlerinden veri toplanmalıdır.

Yanıtlamayanlar tabakasından eşit seçim şansı verilerek ve yerine koymadan tesadüfi seçilen $n_2'=148$ altörnek birimiyle yüz-yüze görüşme yapılmak üzere anketörler Tekel Fabrikasına gitmişlerdir. 148 kişi ile yüz-yüze görüşme yapılmış, yapılan yüz-yüze görüşmeler sonucunda kişilerin birinci soru grubuna ilişkin istatistik değerleri aşağıdaki gibi hesaplanmıştır.

Çizelge 4. 3. Altörnek birimlerine ilişkin istatistik değerleri

Ortalama (\bar{x}_2')	=	2,4155
Mod	=	2
Standart Sapma (s_2')	=	1,2576
Varyans ($s_2'^2$)	=	1,5815

148x21=3108 veri üzerinden yanıtlamayanlar tabakasının ortalaması 2,4155 ve varyansı 1,5815 olarak tahmin edilmiştir. Buradan, çift örnekleme yöntemiyle \bar{X} yığın ortalamasının yansız tahmini hesaplanabilir.

Aşağıda posta anketini yanıtlayan 70 kişi ve posta anketini yanıtlamayan 192 çalışan arasından yüz-yüze görüşme yapılmak üzere tesadüfi olarak seçilen $n_2'=148$ altörnek birimlerine ilişkin istatistik değerleri aşağıdaki gibidir.

Çizelge 4. 4. Posta anketini yanıtlayanlar ve yanıtlamayanlar arasından seçilen altörnek birimlerine ilişkin istatistik değerleri

Yanıtlayanlar	Yanıtlamayanlar
$n_1 = 70$	$n_2 = 192$
$w_1 = 0,27$	$w_2 = 0,73$
$\bar{x}_1 = 3,6960$	$n_2' = 148$
$s_1'^2 = 1,4315$	$\bar{x}_2' = 2,4155$
	$s_2'^2 = 1,5815$

Buradan yığın ortalamasının yansız tahmini,

$$\bar{x}' = \frac{n_1}{n} \bar{x}_1 + \frac{n_2}{n} \bar{x}_2'$$

$$\bar{x}' = \frac{70}{262} 3,6960 + \frac{192}{262} 2,4155$$

$$\bar{x}' = 2,75$$

olarak bulunur. Çift örnekleme yöntemiyle yığın ortalamasının tahminini 2,75 “Kararsızım” olarak yorumlanır.

Yalnızca yanıtlayanların verisine dayalı yığın ortalaması tahmin edilmiş olsaydı, yığın ortalaması 3,6960 olarak bulunacaktı ve yığında bulunan birimlerin birinci soru grubuna verdikleri yanıtların ortalaması “Aynı Fikirdeyim” olarak yorumlanacaktı. Yanıtlamama olması durumu yığın parametrelerine ilişkin tahminlerin yanlı olmasına yol açmaktadır. Birinci soru grubu için yanıtlamama durumundan kaynaklanan farklılık aşağıdaki gibi bulunur.

$$\text{FARKLILIK} = \bar{x}' - \bar{x}_1 = 2,75 - 3,6960 = -0,946$$

Yığın ortalamasının tahminin tek örnekleme yöntemiyle yanlı yapılmış olduğundan, tahminde -0,946 bir farklılık oluşmuştur.

4.1.3.1.3. \bar{x}' 'nin varyans tahmini

Yığın ortalaması tahmini için varyans değeri hesaplanacaktır. Burada σ_2^2 yanıtlamayanlar tabakasının varyansı bilinmediği için yerine yanıtlamayanlar tabakasından alınan altörnekten elde edilen varyans tahmini $s_2'^2=1,5815$ değeri kullanılacaktır. $\sigma^2=\sigma_2^2$ olduğu için σ^2 'nin tahmini de $s^2=s_2'^2=1,5815$ olur.

Örnek ortalamasının varyansı,

$$V(\bar{x}') = \frac{N-n}{(N-1)} \frac{\sigma^2}{n} + \frac{1}{n.N} (k-1) \frac{N_2}{N_2-1} \sigma_2^2$$

formülünden hesaplanır. Burada $\frac{N}{N-1} = 1$ ve $\frac{N_2}{N_2-1} = 1$ olduğu varsayımı yapıldığı için $N=N-1$ ve $N_2=N_2-1$ olarak alındığında \bar{x}' 'nin varyans formülü aşağıdaki gibi olur.

$$V(\bar{x}') = \frac{N-n}{N} \frac{\sigma^2}{n} + \frac{1}{n.N} (k-1) \sigma_2^2$$

Ancak σ_2^2 , σ^2 ve W_2 değerleri bilinmediğinden, bu değerler yerine tahminleri olan $s_2'^2 = 1,5815$, $s^2 = s_2'^2 = 1,5815$ değerleri kullanılacağından varyans formülü tahmine dönüşmüş olur. Örnek ortalamasının varyans tahmini,

$$\begin{aligned} v(\bar{x}') &= \frac{N-n}{N} \frac{s_2'^2}{n} + \frac{1}{n.N} (k-1) s_2'^2 \\ &= \left(\frac{384-262}{384} \frac{1,5815}{262} \right) + \frac{1}{262.384} (1,29-1) 1,5815 \end{aligned}$$

$$v(\bar{x}') = 0,00192$$

olarak bulunmuş olur. Tek örnekleme yöntemiyle ortalamının varyansı 0,0167 olarak bulunurken, çift örnekleme yöntemiyle ortalamının varyansı 0,00192 olarak bulunmuştur. $v(\bar{x}) = 0,0167 > v(\bar{x}') = 0,00192$ olduğu için çift örnekleme yönteminin daha duyarlı bir yöntem olduğu söylenebilir.

4.1.3.2. Cochran'ın çift örnekleme yöntemi

\bar{X} yığının ortalamasının tahmini Hansen ve Hurwitz (9)'in yönteminde olduğu gibi aşağıdaki formülden hesaplanır.

$$\bar{x}' = \frac{n_1}{n} \bar{x}_1 + \frac{n_2}{n} \bar{x}_2'$$

\bar{x}' 'nin bulunabilmesi için yanıtlamayanlar tabakasından alınacak altörnek çapının belirlenerek, altörnek birimlerinden veri toplanılması gerekir.

Cochran (10)'in yöntemiyle yanıtlamama olması durumunda n_2' , n ve k 'nin belirlenmesi iki varsayım altında yapılırken, her iki varsayım altında da \bar{X} yığının ortalamasının tahmini yukarıda verilmiş olan \bar{x}' formülünden hesaplanır.

4.1.3.2.1. $\frac{N}{N-1} = 1$, $\frac{N_2}{N_2-1} = 1$ ve $S^2=S_2'^2$ varsayımı altında Cochran'ın yöntemi

Bu bölümde istenilen tüm hesaplamalar N ve N_2' 'nin yeteri kadar geniş olduğu ve $\frac{N}{N-1} = 1$ ve $\frac{N_2}{N_2-1} = 1$ olduğu varsayımı ile $S^2=S_2'^2$ ele alınan yığının özellikleri ile yanıtlamayanların oluşturduğu yığının özelliklerinin aynı olduğu varsayımı altında yürütülmüştür.

4.1.3.2.1.1. k , n_2' ve n 'in Optimum Değerlerinin Hesaplanması

$n=262$ kişi üzerinden yürütülen anket çalışmasının yalnızca 70'inden yanıt alınabilmektedir. 70 kişiye ilişkin istatistik sonuçları Çizelge 4.2.'de verilmiştir. Yanıtlayanlar Tabakasına Dayalı Hesaplanan İstatistik Değerleri

Buradan istenilen optimum n ve k değerleri hesaplanabilir. Cochran (10)'da optimum altörnekleme kesri k'nın formülü aşağıdaki gibi verilmiştir.

$$k_{\text{opt}} = \sqrt{\frac{c_2(S^2 - W_2 S_2^2)}{S_2^2(c_0 + c_1 W_1)}}$$

$\frac{N}{N-1} = 1$, $\frac{N_2}{N_2-1} = 1$ ve $S^2 = S_2^2$ varsayımı altında istenilen optimum k değeri aşağıdaki formüle dönüşmüş olur.

$$k_{\text{opt}} = \sqrt{\frac{c_2 W_1}{c_0 + c_1 W_1}}$$

Belirlenen $c_0 = 400\ 000$, $c_1 = 1500\ 000$, $c_2 = 5000\ 000$ maliyet kısıtları ile optimum k formülünde W_1 'in tahmini olan $w_1 = 0,27$ değeri yazılacak olursa k_{opt} ,

$$\begin{aligned} k_{\text{opt}} &= \sqrt{\frac{c_2 W_1}{c_0 + c_1 W_1}} \\ &= \sqrt{\frac{5000000 \times 0,27}{400000 + 1500000 \times 0,27}} \end{aligned}$$

$$k_{\text{opt}} = 1,29$$

şeklinde hesaplanır. Buradan yanıtlamayanlardan alınacak altörnek genişliği n_2' ,

$$n_2' = n_2 / k_{\text{OPT}}$$

$$n_2' = 192 / 1,29 = 148$$

bulunur. Yanıtlamayanlar tabakasından alınacak altörnek genişliği Hansen ve Hurwitz'in yönteminde olduğu gibi 148 olarak bulunmuştur. 148 kişiden veri toplamak amacıyla yüz-yüze görüşme anketi uygulanacağına karar verilmiş oldu.

Yanıtlamama olması durumunda yığından seçilecek ilk örnek genişliği,

$$n_{opt} = \frac{N[S^2 + (k_{opt} - 1)W_2 S_2^2]}{NV + S^2}$$

formülünden hesaplanır. n_{opt} formülünde V varyans aşağıdaki gibidir.

$$V = V(\bar{x}') = \frac{N-n}{N} \frac{S^2}{n} + \frac{(k-1) W_2 S_2^2}{n}$$

V varyans eşitliği n_{opt} 'da yerine yazılacak olursa optimum n aşağıdaki gibi olur.

$$n_{opt} = \frac{N[S^2 + (k_{opt} - 1)W_2 S_2^2]}{N \left\{ \frac{N-n}{N} \frac{S^2}{n} + \frac{(k_{opt} - 1) W_2 S_2^2}{n} \right\} + S^2}$$

$S^2 = S_2^2$ varsayımından n_{opt} 'da pay ve payda S^2 ortak parantezine alınacak olursa,

$$n_{opt} = \frac{S^2 N [1 + (k_{opt} - 1)W_2]}{S^2 N \left\{ \left(\frac{N-n}{N} \frac{1}{n} \right) + \frac{(k_{opt} - 1) W_2}{n} \right\} + S^2}$$

formülde S^2 'ler sadeleştirilebilir. Buradan n_{opt} ,

$$n_{opt} = \frac{N [1 + (k_{opt} - 1)W_2]}{N \left\{ \left(\frac{N-n}{N} \frac{1}{n} \right) + \frac{(k_{opt} - 1) W_2}{n} \right\} + 1}$$

formülüne dönüşmüş olur. n_{opt} formülünde, bulmuş olduğumuz optimum k değerini, W_2 yerine tahmini olan $w_2=0,73$ değerini, n ve N derlerini yerine yazılırsa,

$$n_{opt} = \frac{384[1 + (1,29 - 1)0,73]}{384 \left\{ \left(\frac{384 - 262}{384} \frac{1}{262} \right) + \frac{(1,29 - 1) 0,73}{262} \right\} + 1}$$

$$n_{opt} = 262$$

yanıtlamama olması durumunda yığından alınması gereken optimum örnek çapı 262 olarak bulunur. Tek örnekleme yöntemiyle yığından alınacak optimum örnek çapı ile Cochran'ın çift örnekleme yöntemiyle yanıtlamama olması durumunda yığından alınacak optimum ilk örnek çapı aynı olup 262 olarak bulunmuştur. Hansen ve Hurwitz yönteminde ise bu örnek çapı 317 olarak belirlenmişti.

4.1.3.2.1.2. Yığın ortalamasının tahmini

Yanıtlamama olması durumu için çift örnekleme yöntemiyle yığın ortalamasının tahmini Hansen ve Hurwitz'in yönteminde olduğu gibi aşağıdaki formülle hesaplanır.

$$\bar{x}' = \frac{n_1}{n} \bar{x}_1 + \frac{n_2}{n} \bar{x}_2'$$

Hansen ve Hurwitz'in yönteminde olduğu gibi, $\frac{N}{N-1} = 1$, $\frac{N_2}{N_2-1} = 1$ ve $S^2 = S_2'^2$ varsayımı altında Cochran'ın yönteminde de yanıtlamayanlar arasından yüz-yüze görüşme yapılmak üzere seçilen altörnek birimlerinin sayısı $n_2' = 148$ olduğu için, bu kişilere ilişkin istatistik değerleri aynı olup Çizelge 4.4.'teki gibidir.

Buradan Çizelge 4.4.'teki istatistik değerlerinden yığın ortalamasının yansız tahmini,

$$\bar{x}' = \frac{n_1}{n} \bar{x}_1 + \frac{n_2}{n} \bar{x}_2'$$

$$\bar{x}' = \frac{70}{262} 3,6960 + \frac{192}{262} 2,4155$$

$$\bar{x}' = 2,75$$

Hansen ve Hurwitz'de olduğu gibi 2,75 olarak bulunmuş olur.

4.1.3.2.1.3. \bar{x}' 'nün varyans tahmini

Örnek ortalamasının varyansı aşağıdaki formülden hesaplanır.

$$V(\bar{x}') = \frac{N-n}{N} \frac{S^2}{n} + \frac{(k-1) W_2 S_2^2}{n}$$

$S^2=S_2^2$ varsayımı altında S^2 yığın varyansının S_2^2 yanıtlamayanlar tabakasının varyansına eşit olduğu kabul edilmiş olur. $V(\bar{x}')$ formülünde S_2^2 ve W_2 değerleri bilinmediğinden yerlerine tahminleri olan $s_2'^2 = 1,5815$, $w_2=0,73$ değeri ve bilinmeyen S^2 yığın varyansı için, $S^2=S_2^2$ varsayımı altında, $s^2 = s_2'^2 = 1,5815$ tahmini kullanılır. Böylece $V(\bar{x}')$ formülü tahmine dönüşmüş olur;

$$v(\bar{x}') = \frac{N-n}{N} \frac{s^2}{n} + \frac{(k-1) w_2 s_2'^2}{n}$$

$$= \frac{384-262}{384} \frac{1,5815}{262} + \frac{(1,29-1) 0,73 (1,5815)}{262}$$

$$v(\bar{x}') = 0,00319$$

Tek örnekleme yöntemiyle ortalamanın varyans tahmini 0,0167, Hansen ve Hurwitz'in çift örnekleme yöntemiyle ortalamanın varyans tahmini 0,00192 olarak bulunurken, Cochran'da bu değer 0,00319 olarak bulunmuştur.

\bar{x}' örnek ortalamasının varyans tahmini ikinci bir yolla, Cochran (10) ve Rao (46) çıkarsamalarından bulunan \bar{x}' 'nin varyans tahmini formülü olan;

$$v(\bar{x}') = \frac{(N-n)(n_1-1)}{N(n-1)} w_1 \frac{s_1^2}{n_1} + \frac{(N-1)(n_2-1) - (n-1)(n_2'-1)}{N(n-1)} w_2 \frac{s_2'^2}{n_2'} + \frac{N-n}{N(n-1)} [w_1(\bar{x}_1 - \bar{x}')^2 + w_2(\bar{x}_2' - \bar{x}')^2]$$

s_1^2 : n_1 birimlerinin varyansı

$s_2'^2$: n_2' birimlerinin varyansı

kullanarak hesaplanabilir. $v(\bar{x}')$ formülünde Çizelge 4.1.6'da verilen değerler yerine yazıldığında örnek ortalamasının varyansı,

$$v(\bar{x}') =$$

$$\frac{(384-262)(70-1)}{384(262-1)} 0,27 \frac{1,4315}{70} + \frac{(384-1)(192-1) - (262-1)(148-1)}{384(262-1)} 0,73 \frac{1,5815}{148}$$

$$+ \frac{384-262}{384(262-1)} [0,27(3,6960 - 2,75)^2 + 0,73(2,4155 - 2,75)^2]$$

$$v(\bar{x}') = 0,00356$$

olarak bulunur. Cochran'da her iki varyans formülüyle elde edilen sonuçların hemen aynı olduğu görülmektedir.

4.1.3.2.2. $\frac{N}{N-1} \neq 1, \frac{N_2}{N_2-1} \neq 1$ ve $S^2 \neq S_2^2$ varsayımı altında Cochran'ın yöntemi

$\frac{N}{N-1} = 1$ ve $\frac{N_2}{N_2-1} = 1$ varsayımın yapılmadığı ve yığın varyansı ile yanıtlamayanlar tabakasının varyansının eşit kabul edilmediği durumda, yanıtlamama olması durumunda \bar{X} yığın ortalamasının yansız tahmini yapılmaya çalışılmıştır.

\bar{X} yığın ortalamasının tahmini aşağıdaki formülden hesaplanır.

$$\bar{x}' = \frac{n_1}{n} \bar{x}_1 + \frac{n_2}{n} \bar{x}_2'$$

\bar{x}' 'nin bulunabilmesi için yanıtlamayanlar tabakasından alınacak altörnek çapının belirlenerek, altörnek birimleriyle yüz-yüze görüşme yapılarak veri toplanması gerekir.

4.1.3.2.2.1. k, n_2' ve n 'in optimum değerlerinin hesaplanması

Araştırma yalnızca posta anketi yöntemine dayalı yürütülmüş (tek örnekleme yöntemi yapılmış) olsaydı yığın varyansı (S^2) tahmininin yanıtlayanların varyansına eşit olduğu ve S^2 yığın varyansının tahmini olan s^2 'nin de $s_1^2=1,4315$ eşit olduğu kabul edilecekti.

Yanıtlamama olması durumunda, $S^2 \neq S_2^2$ varsayımı altında, yığından seçilecek optimum örnek genişliği,

$$n_{opt} = \frac{N[S^2 + (k_{opt} - 1)W_2 S_2^2]}{NV + S^2}$$

formülünden hesaplanırken, optimum altörnekleme kesri,

$$k_{\text{opt}} = \sqrt{\frac{c_2(S^2 - W_2 S_2^2)}{S_2^2(c_0 + c_1 W_1)}}$$

formülünden hesaplanır. n_{opt} ve k_{opt} formüllerinde S^2 , S_2^2 ve V bilinmeyenlerdir.

$S^2 \neq S_2^2$ varsayımı altında n ve k 'nin optimum değerlerini hesaplayabilmek için yanıtlayanlar ve yanıtlamayanlar tabakalarına ilişkin varyansların biliniyor olması istenir. Yanıtlayanlar tabakasının varyans tahmini $s_1^2=1,4315$ dir, ancak yanıtlamayanlar tabakasına ilişkin herhangi bir bilgi mevcut olmadığından varyans değeri (S_2^2) de bilinmemektedir.

S_2^2 'yi tahmin edebilmek amacıyla yanıtlamayanlar arasından tesadüfi olarak 30 çaplı ön örnek seçilmesine karar verilmiştir. Öntest çalışması yanıtlamayan birimleriyle yüz-yüze görüşme yapılması şeklinde olmuş ve elde edilen bilgiler Çizelge 4.5.'te verilmiştir.

Çizelge 4. 5. Posta anketini yanıtlayanlara ve yanıtlamayanlar arasından seçilen $n_{2\delta}=30$ ön örnek birimlerine ilişkin istatistik değerleri

Yanıtlayanlar	Yanıtlamayanlar
$n_1=70$	$n_2=192$
$w_1=0,27$	$w_2=0,73$
$\bar{x}_1=3,6960$	$n_{2\delta}=30$
$s_1^2=1,4315$	$\bar{x}_{2\delta}=2,3721$
	$s_{2\delta}^2=1,1528$

- $n_{2\delta}$: Yanıtlamayanlardan seçilen öntest birimlerinin sayısı
 $\bar{x}_{2\delta}$: Yanıtlamayanlara ilişkin öntestten hesaplanan ortalama tahmini
 $s_{2\delta}^2$: Yanıtlamayanlara ilişkin öntestten hesaplanan varyans tahmini

S^2 yığın varyansının tahmini, posta anketini yanıtlayan $n=70$ kişinin verisine dayalı olarak $s_1^2=1,4315$ değerine eşit kabul edilmiştir. S_2^2 yığın varyansının tahmini, yanıtlamayanlardan seçilen öntest çalışması sonuçlarına dayalı olarak $s_{20}^2= 1,1528$ olarak bulunmuştur. Buradan $\frac{N}{N-1} \neq 1$, $\frac{N_2}{N_2-1} \neq 1$ ve $S^2 \neq S_2^2$ varsayımı altında optimum n ve k 'yi hesaplanabilir.

Belirlenen maliyet kısıtları altında optimum k değeri,

$$k_{opt} = \sqrt{\frac{c_2(S^2 - W_2 S_2^2)}{S_2^2(c_0 + c_1 W_1)}}$$

formülünden hesaplanır. Optimum k 'yi,

$$k_{opt} = \sqrt{\frac{5000000 (1,4315 - (0,73) \cdot 1,1528)}{1,1528(400000 + 1500000 \cdot (0,27))}}$$

$$k_{opt} = 1,78$$

olarak bulunur. Buradan yanıtlamayanlar tabakasından yüz-yüze görüşme yapılmak üzere seçilecek altörnek genişliği n_2' 'ü,

$$n_2' = n_2 / k_{OPT}$$

$$n_2' = 192 / 1,78 = 107$$

hesaplanmış olur.

Yanıtlamama olması durumunda yığından seçilecek optimum örnek genişliği,

$$n_{opt} = \frac{N[S^2 + (k_{opt} - 1)W_2 S_2^2]}{NV + S^2}$$

formülünden bulunur. n_{opt} denkleminde $V = V(\bar{x}')$ 'in belirlenmesi gerekir. V varyans formülünde; S^2 , S_2^2 ve W_2 yığın parametreleri bilinmediği için, parametre değerleri yerine sırasıyla $s_1^2=1,4315$, $s_2^2= 1,1528$ ve $w_2=0,73$ tahmin değerleri kullanıldığında, $V(\bar{x}')$ 'in tahmini yapılmış olur. $V(\bar{x}')$ 'nin tahmini,

$$\begin{aligned} v(\bar{x}') &= \frac{N-n}{n} \frac{s^2}{n} + \frac{(k-1) w_2 s_2'^2}{n} \\ &= \frac{384-262}{384} \frac{1,4315}{262} + \frac{(1,78-1) 0,73 1,1528}{262} \\ &= 0,00424 \end{aligned}$$

olarak bulunur. n_{opt} formülünde tüm değerler yerine yazıldığında optimum ilk örnek genişliği,

$$n_{opt} = \frac{384[1,4315 + (1,78-1) 0,73 1,1528]}{384 (0,00424) + 1,4315}$$

$$n_{opt} = 262$$

bulunmuş olur. Optimum örnek genişliği, Cochran'ın yönteminde $\frac{N}{N-1} = 1$ ve

$\frac{N_2}{N_2-1} = 1$ ile $S^2 = S_2^2$ olduğu varsayımının yapıldığı durum ve yapılmadığı durum

için aynı olup 262 olarak bulunmuştur.

4.1.3.2.2.2. Yığın ortalamasının tahmini

$\frac{N}{N-1} \neq 1$, $\frac{N_2}{N_2-1} \neq 1$ ve $S^2 \neq S_2^2$ varsayımı altında yanıtlanmayanlar tabakasından

seçilecek altörnek genişliği $n_2' = 107$ olarak hesaplanmıştır. Yanıtlanmayanlar

arasından tesadüfi olarak seçilen 107 kişiyle yüz-yüze görüşme yapılmış ve bu kişilere ilişkin verilere dayalı olarak hesaplanan istatistik değerleri aşağıdaki gibi bulunmuştur.

Çizelge 4. 6. Altörnek birimlerine ilişkin istatistik değerleri

Ortalama (\bar{x}_2')	=	2,5581
Mod	=	2
Standart Sapma (s_2')	=	1,2788
Varyans ($s_2'^2$)	=	1,6354

107x21=2247 veri üzerinden yanıtlamayanlar tabakasının ortalaması 2,5581 ve varyansı 1,5815 olarak tahmin edilmiştir. Buradan, çift örnekleme yöntemiyle \bar{X} yığın ortalamasının yansız tahmini hesaplanabilir.

Yığın ortalamasının çift örnekleme yöntemiyle tahmini Çizelge 4.7.'de verilen değerlerin \bar{x}' formülünde yerine yazılmasıyla bulunur.

Çizelge 4. 7. Posta anketini yanıtlayanlar ve yanıtlamayanlar arasından seçilen altörnek birimlerine ilişkin istatistik değerleri

Yanıtlayanlar	Yanıtlamayanlar
$n_1 = 70$	$n_2 = 192$
$w_1 = 0,27$	$w_2 = 0,73$
$\bar{x}_1 = 3,6960$	$n_2' = 107$
$s_1'^2 = 1,4315$	$\bar{x}_2' = 2,5581$
	$s_2'^2 = 1,6354$

Yanıtlamama olması durumunda yığın ortalamasının çift örnekleme yöntemiyle tahmini,

$$\bar{x}' = \frac{n_1}{n} \bar{x}_1 + \frac{n_2}{n} \bar{x}_2'$$

$$\bar{x}' = \frac{70}{262} 3,6960 + \frac{192}{262} 2,5581$$

$$\bar{x}' = 2,86$$

2,86'dır. Birinci soru grubu için yığın ortalamasının tahmini 2,86 "Kararsızım" olarak bulunmuştur.

4.1.3.2.2. 3. \bar{x}' 'nün varyans tahmini

Cochran'ın yöntemiyle örnek ortalamasının varyansı $\frac{N}{N-1} = 1$, $\frac{N_2}{N_2-1} = 1$ ve $S^2 = S_2'^2$ varsayımında olduğu gibi iki formülle hesaplanabilir. \bar{x}' 'nün varyansı,

$$V(\bar{x}') = \frac{N-n}{N} \frac{S^2}{n} + \frac{(k-1) W_2 S_2'^2}{n}$$

dir. V varyans formülünde; S^2 , $S_2'^2$ ve W_2 yığın parametreleri bilinmediği için, parametre değerleri yerine sırasıyla $s_1^2 = 1,4315$, $s_2'^2 = 1,1528$ ve $w_2 = 0,73$ tahmin değerleri kullanıldığında, $V(\bar{x}')$ 'in tahmini aşağıdaki gibi hesaplanır.

$$\begin{aligned} v(\bar{x}') &= \frac{N-n}{N} \frac{s^2}{n} + \frac{(k-1) w_2 s_2'^2}{n} \\ &= \frac{384-262}{384} \frac{1,4315}{262} + \frac{(1,78-1) 0,73 1,1528}{262} \\ &= 0,00424 \end{aligned}$$

Cochran (10) ve Rao (47) çıkarsamalarından varyans tahmini,

$$v(\bar{x}') = \frac{(N-n)(n_1-1)}{N(n-1)} w_1 \frac{s_1^2}{n_1} + \frac{(N-1)(n_2-1) - (n-1)(n_2'-1)}{N(n-1)} w_2 \frac{s_2'^2}{n_2'}$$

$$\begin{aligned}
& + \frac{N-n}{N(n-1)} [w_1(\bar{x}_1 - \bar{x}')^2 + w_2(\bar{x}_2' - \bar{x}')^2] \\
& = \frac{(384-262)(70-1)}{384(262-1)} 0,27 \frac{1,4315}{70} + \frac{(384-1)(192-1) - (262-1)(107-1)}{384(262-1)} 0,73 \frac{1,6354}{107} \\
& + \frac{384-262}{384(262-1)} [0,27(3,6960 - 2,86)^2 + 0,73(2,5581 - 2,86)^2]
\end{aligned}$$

$$v(\bar{x}') = 0,00583$$

0,00583 olarak hesaplanır. $\frac{N}{N-1} \neq 1$, $\frac{N_2}{N_2-1} \neq 1$ ve $S^2 \neq S_2^2$ varsayımı altında

Cochran'ın yöntemiyle hesaplanan ortalamanın varyans tahmini 0,00424 ve 0,00583 olarak bulunmuştur. Varyans değerleri birbirine yakın bulunmuştur.

4.1.3.3. Srinath'ın çift örnekleme yöntemi

Srinath'ın çift örnekleme yöntemiyle \bar{X} yığın ortalaması tahmini Hansen ve Hurwitz (9) formülüyle verilmiştir.

$$\bar{x}' = \frac{n_1}{n} \bar{x}_1 + \frac{n_2}{n} \bar{x}_2'$$

Buna karşın \bar{x}' için varyans formülü,

$$V(\bar{x}') = \left(\frac{N-n}{Nn}\right)S^2 + \frac{k'S_2^2}{n}$$

şeklinde verilmiştir.

Yanıtlamama olması durumunda, \bar{X} yığın ortalamasının yansız bir tahmininin yapılabilmesi için öncelikle, yanıtlamayanlar tabakasından alınacak optimum altörnek genişliğinin hesaplanması gerekir.

4.1.3.3.1. k , n_2' ve n' 'in optimum değerlerinin hesaplanması

Bu bölümde yanıtlamama olması durumunda yığından seçilmesi gereken örnek genişliği n ve altörnek genişliği n_2' hesaplanmaya çalışılacaktır. Bunun için öncelikle, Srinath'ın önermiş olduğu, k' altörnek genişliği aşağıdaki formülden hesaplanmaya çalışılacaktır.

$$k'_{opt} = \sqrt{\frac{(S^2 - W_2 S_2^2) c_2 W_2^2}{S_2^2 (c_0 + c_1 W_1)}} - W_2 \quad [44]$$

k'_{opt} formülü, $\frac{N}{N-1} \neq 1$, $\frac{N_2}{N_2-1} \neq 1$ ve $S^2 \neq S_2^2$ varsayımı altında Cochran'ın yöntemiyle hesaplanan k_{opt} 'a göre tanımlanacak olursa, aşağıdaki gibi yazılabilir.

$$k'_{opt} = (k_{opt} - 1) W_2$$

$k_{opt} = 1,78$ olarak bulunmuştu. Buradan, optimum k' değeri,

$$k'_{opt} = (k_{opt} - 1) W_2$$

$$k'_{opt} = (1,78-1).0,73$$

$$k'_{opt} = 0,59$$

şeklinde hesaplanmış olur. $k'_{opt} < 1$ olduğu için altörnekleme yapılmayacağı, yanıtlamayan tüm birimleriyle yüz-yüze görüşme yapılacağı sonucuna ulaşılmıştır. Tüm birimlerle yüz-yüze görüşme yapılması zaman alacağından ve daha yüksek maliyet gerektireceğinden uygulaması yapılmamıştır. -

Aşağıda yer alan Çizelge 4.8.'de birinci soru grubuna ilişkin tek örnekleme ve çift örnekleme yöntemleriyle elde edilmiş olan sonuç istatistikleri çizelge halinde verilmiştir.

Çizelge 4. 8. Birinci soru grubu sonuç istatistikleri

YÖNTEMLER	Varsayımlar	k_{opt}	n_2'	n	Ortalama tahmini	Ortalama tahmininin varyans tahmini	Cochran (10) ve Rao (46) çıkarsamalarından Ortalamanın Varyans tahmini
Tek örnekleme Yöntemi	-	-	-	262	3,6960	0,0167	-
Hansen ve Hurwitz'in Çift Örnekleme Yöntemi	$\frac{N}{N-1} = 1, \frac{N_2}{N_2-1} = 1$ ve $\sigma^2 = \sigma_2^2$ varsayımı	1,29	148	317	2,75	0,00192	-
Cochran'ın Çift Örnekleme Yöntemi	$\frac{N}{N-1} = 1, \frac{N_2}{N_2-1} = 1$ ve $S^2 = S_2^2$ varsayımı	1,29	148	262	2,75	0,00319	0,00356
	$\frac{N}{N-1} \neq 1, \frac{N_2}{N_2-1} \neq 1$ ve $S^2 \neq S_2^2$ varsayımı	1,78	107	262	2,86	0,00424	0,00583
Srinath'ın Çift	-	0,59	-	-	-	-	-

Burada $p=0,27, q=0,73$ dir.

Çizelgeye bakıldığında, yığın ortalamasının tahmini tek örnekleme yöntemiyle 3,696

olurken, Hansen ve Hurwitz'in ve $\frac{N}{N-1} = 1, \frac{N_2}{N_2-1} = 1$ ve $S^2 = S_2^2$ varsayımı

altında Cochran'ın çift örnekleme yöntemleriyle 2,75 olarak, $\frac{N}{N-1} \neq 1, \frac{N_2}{N_2-1} \neq 1$

ve $S^2 \neq S_2^2$ varsayımı altında ise 2,86 olarak bulunmuştur. Srinath'ın çift örnekleme yöntemi $k_{opt}=0,59$ bulunduğu için kullanılamamıştır. Birinci soru grubu, işyerinde

çalışanlar üzerinde örgütsel adaletin uygulanıp uygulanmadığını ölçen sorulardan oluşmaktadır. Dolayısıyla, tek örnekleme yöntemine dayalı olarak; çalışanların örgütsel adaletin uygulandığı görüşünde oldukları, çift örnekleme yöntemine dayalı olarak ise; örgütsel adaletin uygulandığı konusunda çalışanların kararsız oldukları tespit edilmiştir.

Çizelge 4.8.'de ortalama tahminin varyans tahminleri kolonuna bakıldığında, en yüksek varyans değerinin tek örnekleme yöntemine ait olduğu ve bu değer 0,0167 olduğu görülmektedir. Tek örnekleme yönteminin varyansı en büyük olduğu için, çift örnekleme yöntemlerinin tek örnekleme yönteminden daha duyarlı olduğu söylenebilir. Örnekleme yöntemleri arasında en düşük varyanslı olanı $\frac{N}{N-1} = 1$, $\frac{N_2}{N_2-1} = 1$ ve $\sigma^2 = \sigma_2^2$ varsayımıyla Hansen ve Hurwitz'in çift örnekleme yöntemi olduğu için, en duyarlı örnekleme yöntemi Hansen ve Hurwitz'in çift örnekleme yöntemidir denir. Ancak, çizelgeye bakıldığında, tüm çift örnekleme yöntemleriyle bulunan varyans tahmini değerleri arasında çok da büyük farklılıklar olmadığı görülmektedir.

n_2' altörnek çaplarına bakıldığında, en düşük altörnek çapının Cochran'ın $\frac{N}{N-1} \neq 1$,

$\frac{N_2}{N_2-1} \neq 1$ ve $S^2 \neq S_2^2$ varsayımı altında hesaplanan altörnek çapına ait olduğu

görülmektedir. Bu yöntemle $n_2'=107$ olarak bulunmuş ve posta ve yüz-yüze görüşme anketi yöntemleriyle toplam görüşme yapılan kişi sayısı 177 kişi olmuştur. Diğer çift örnekleme yöntemlerinde ise altörnek çapı 148, toplam görüşme yapılan kişi sayısı ise 218 kişi olmuştur. Buradan, çift örnekleme yöntemleri arasında Cochran'ın

$\frac{N}{N-1} \neq 1$, $\frac{N_2}{N_2-1} \neq 1$ ve $S^2 \neq S_2^2$ varsayımı altında kullanılan yöntemin daha düşük

maliyet gerektirdiği görülmüştür.

4.2. İkinci Soru Grubu

İkinci soru grubu, işletme çalışanlarının kendilerini örgüte bağlı hissedip hissetmediklerini ölçen sorulardan oluşmaktadır. Birinci soru grubunda olduğu gibi, yığına ilişkin istenilen ortalama tahmini tek örnekleme ve Hansen ve Hurwitz'in ile Cochran'ın çift örnekleme yöntemleri ile yapılmaya çalışılmıştır.

70 kişiye yöneltilen ikinci soru grubunda 17 soru vardır.

4.2.1. n örnek çapının belirlenmesi

İkinci soru grubuna göre, $N=384$ çaplı yığından seçilecek örnek çapını belirlenmeye çalışılacaktır. En uygun örnek büyüklüğü $d = z \cdot sh(\bar{x})$ formülünden hesaplanmaya çalışılmıştır.

Örnek çapının tahmininde, S^2 yığında birim başına düşen varyans bilinmediğinden, yığından $n_0 = 40$ çapında bir ön örnek seçilmiştir.

Anket formları, yığından tesadüfi seçilen ön örnek birimi 40 kişiye postalanmıştır. 40 kişi üzerinden yapılan öntest çalışması sonucunda aşağıdaki çizelgede verilen istatistiksel değerlere ulaşılmıştır.

Çizelge 4. 9. Öntest sonucunda elde edilen istatistik değerleri

Ortalama (\bar{x})	=	4,1062
Ortalamanın standart hatası ($sh(\bar{x})$)	=	0,1762
Mod	=	5
Standart Sapma (s)	=	1,1773
Varyans (s^2)	=	1,3860

Öntest çalışmasıyla, $40 \times 17 = 680$ veri üzerinden yığın varyansı $s^2 = 1,3860$ olarak tahmin edilmiştir.

$d = z \cdot sh(\bar{x})$ formülünde 0,05 hata düzeyinde z değeri 2, $d=0,09$ olarak alındığında yığından seçilecek örnek çapı,

$$d = z \cdot \sqrt{\frac{N-n}{N} \frac{S^2}{n}}$$

$$(0,09)^2 = 4 \cdot \frac{384-n}{384} \frac{1,386}{n}$$

$$n = 246$$

şeklinde bulunur.

4.2.2. Tek örnekleme yöntemi

Öntest çalışmasıyla yığından tesadüfi seçilen 40 kişiyle görüşme yapılmıştı dolayısıyla anket çalışması yığından tesadüfi seçilecek $246-40=206$ kişiye uygulanacaktır. Yığından tesadüfi seçilen 206 kişinin adreslerine anket formları postalanmıştır. Postalanan anket formlarını yanıtlayanların sayısı $40+30 = 70$ kişi olmuştur. Tüm hatırlatma ve girişimlere rağmen posta anketini yanıtlamayan 176 kişidir.

Tek örnekleme yöntemiyle yığın ortalamasının tahmini $n= 246$ kişi üzerinden değil de, yalnızca posta anketini yanıtlayan 70 kişi üzerinden elde edilen verilere dayalı olarak yapılır.

70 kişiye ilişkin verilerden, yığın parametre tahminleri aşağıdaki çizelgedeki gibidir.

Çizelge 4. 10. Yanıtlayanlar tabakasına dayalı hesaplanan ıstatistik değeri

<i>Ortalama (\bar{x})</i>	=	3,9013
<i>Ortalamanın standart hatası ($sh(\bar{x})$)</i>	=	0,1286
<i>Mod</i>	=	5
<i>Standart Sapma (s)</i>	=	1,1903
<i>Varyans (s^2)</i>	=	1,4168

4.2.2.1. Yiğın ortalamasının tahmin edilmesi

70 kişiye yöneltilen ikinci soru grubunda 17 soru vardır. Toplam $70 \times 17 = 1190$ veri üzerinden yiğın ortalamasının tahmininin $\bar{x} = 3,9$ "Aynı Fikirdeyim" olarak bulunmuştur.

Buradan, kişilerin genelinin çalıştıkları kuruma bağlılıklarının yüksek olduğu, söylenebilir.

4.2.2.2. \bar{x} 'nin varyans tahmini

Yiğın varyansının tahmini Çizelge 4.10.'da $s^2 = 1,4168$ olarak bulunmuştur. Örnek ortalamasının varyansı yalnızca yanıtlayanlara bağlı olarak hesaplanacak olursa, yanıtlamama olması durumu göz ardı edilmiş olacaktır ve n örnek çapı 246 olması gerekirken, 70 olduğu kabul edilecektir. İstenilen varyans değeri aşağıdaki formülden hesaplanmış olacaktır.

$$V(\bar{x}) = \frac{N-n}{N} \frac{S^2}{n}$$

Burada $N=384$, örnek çapı $n=70$ ve S^2 yiğın varyansının tahmini $s^2=1,4168$ olarak alındığında, örnek ortalamasının varyans tahmini aşağıdaki gibi bulunur.

$$v(\bar{x}) = \frac{384-70}{384} \frac{1,4168}{70} = 0,01655$$

4.2.3. Yanıtlamama için çift örnekleme yöntemi

İkinci soru grubu üzerinden yürütülen çalışmada, posta anketini yanıtlayanların oranı %28 olmuştur. Burada da yanıtlamama için Hansen ve Hurwitz'in, Cochran'ın ve Srinath'ın çift örnekleme yöntemi kullanılarak istenilen parametre tahminleri ikinci soru grubu için yapılmaya çalışılmıştır.

4.2.3.1. Hansen ve Hurwitz'in yöntemi

Bu bölümde istenilen tüm hesaplamalar $\frac{N}{N-1} = 1$ ve $\frac{N_2}{N_2-1} = 1$ olduğu varsayımı ile

$\sigma^2 = \sigma_2^2$ varsayımı altında yürütülmüştür.

4.2.3.1.1. k, n₂' ve n'in optimum değerlerinin hesaplanması

$\frac{N}{N-1} = 1$, $\frac{N_2}{N_2-1} = 1$ ve $\sigma^2 = \sigma_2^2$ varsayımı altında yanıtlamama olması durumunda optimum n, n₂' ve k hesaplanmaya çalışılacaktır.

Yanıtlamama olması durumunda optimum n örnek çapı,

$$n_{OPT} = \hat{n} [1 + (k-1)q]$$

formülünden hesaplanır.

n=246 kişi üzerinden yürütülen anket çalışmasının yalnızca 70'inden yanıt alınabilmektedir. 176 kişiden yanıt alınmamıştır.

Yanıtlayanlar	Yanıtlamayanlar
$n_1 = 70$	$n_2 = 176$
$p = \frac{n_1}{n} = \frac{70}{344} = 0,28 = w_1$	$q = \frac{n_2}{n} = \frac{176}{344} = 0,72 = w_2$

Buradan, $p=0,28$, $q=0,72$ olarak belirlenmiştir. n_{OPT} bulabilmek için için optimum altörnekleme kesri k 'nın hesaplanması gerekir.

Belirlenen $c_0=400\ 000$, $c_1=1500\ 000$, $c_2=5000\ 000$ maliyet kısıtları ile p ve q değerleri altında optimum altörnekleme kesri,

$$k_{OPT} = \sqrt{\frac{c_2 p}{c_0 + c_1 p}} = \sqrt{\frac{5000000 \times 0,28}{400000 + 1500000 \times 0,28}}$$

$$k_{opt}=1,3$$

1,3 olarak bulunur. Buradan yanıtlamayanlar tabakasından alınacak altörnek çapı n_2' 'ü aşağıdaki gibi hesaplanır.

$$n_2' = n_2 / k_{OPT}$$

$$n_2' = 176 / 1,3 = 135$$

Yanıtlamayanlar tabakasından alınacak altörnek çapı 135'tir. 135 kişiden veri toplamak amacıyla yüz-yüze görüşme anketi uygulanmalıdır.

Yanıtlamama olması durumu için yığından seçilmesi gereken optimum örnek çapı aşağıdaki gibi bulunur.

$$n_{opt} = 246 \{1 + (1,3 - 1)0,72\}$$

$$n_{opt} = 299$$

Yığından alınacak örnek çapı 246 olarak belirlenmişti. Burada ise yanıtlamama olması durumunda yığından seçilecek optimum örnek genişliği 299 olarak saptanmıştır.

4.2.3.1.2. Yığın ortalamasının tahmin edilmesi

Yanıtlamama olması durumunda çift örnekleme yöntemiyle yığın ortalamasının yansız tahmini aşağıdaki gibi olduğu verilmiştir.

$$\bar{x}' = \frac{n_1}{n} \bar{x}_1 + \frac{n_2}{n} \bar{x}_2'$$

\bar{x}' formülünde \bar{x}_2' bilinmemektedir. \bar{x}_2' 'ü altörnek ortalamasını bulabilmek için, yanıtlamayanlar arasından seçilecek $n_2'=135$ altörnek birimlerinden veri toplanmalıdır.

Yanıtlamayanlar tabakasından tesadüfi olarak seçilen $n_2'=135$ altörnek birimiyle anketörler yüz-yüze görüşme yapılmak üzere Tekel Fabrikasına gitmişlerdir. Yapılan görüşmeler sonucunda kişilerin ikinci soru grubuna ilişkin istatistik değerleri aşağıdaki gibi elde edilmiştir.

Çizelge 4. 11. Altörnek birimlerine ilişkin istatistik değerleri

Ortalama (\bar{x}_2')	=	2,6686
Mod	=	2,00
Standart Sapma (s_2')	=	1,2987
Varyans ($s_2'^2$)	=	1,6865

135x17=2295 veri üzerinden yanıtlamayanlar tabakasının ortalaması $\bar{x}_2'=2,6686$ ve varyansı $s_2'^2=1,6865$ olarak bulunmuştur.

Aşağıda posta anketini yanıtlayan 70 kişi ve posta anketini yanıtlanmayan 176 çalışan arasından yüz-yüze görüşme yapılmak üzere tesadüfi olarak seçilen $n_2'=135$ altörnek birimlerine ilişkin istatistik değerleri aşağıdaki Çizelge 4.12'deki gibidir.

Çizelge 4. 12. Posta anketini yanıtlayanlar ve yanıtlanmayanlar arasından seçilen altörnek birimlerine ilişkin istatistik değerleri

<i>Yanıtlayanlar</i>	<i>Yanıtlanmayanlar</i>
$n_1 = 70$	$n_2 = 176$
$w_1 = 0,28$	$w_2 = 0,72$
$\bar{x}_1 = 3,9013$	$n_2' = 135$
$s_1^2 = 1,4168$	$\bar{x}_2' = 2,6686$
	$s_2^2 = 1,6865$

Buradan yığın ortalamasının yansız tahmini aşağıdaki gibi hesaplanmış olur.

$$\bar{x}' = \frac{70}{246} 3,9013 + \frac{176}{246} 2,6686$$

$$\bar{x}' = 3,01$$

Çift örnekleme yöntemiyle ikinci soru grubu için yığın ortalamasının tahminini 3,01 “Kararsızım” olarak yorumlanır.

Yalnızca yanıtlayanların verisine dayalı olarak, yığın ortalaması tahmin edilmiş olsaydı, yığın ortalaması 3,9013 olup “Aynı fikirdeyim” olarak yorumlanacaktır.

Yanıtlanmama durumu yığın parametrelerine ilişkin tahminlerin yanlı olmasına yol açmaktadır. İkinci soru grubu için yanıtlanmama durumundan kaynaklanan farklılık aşağıdaki gibi bulunur.

$$\text{FARKLILIK} = \bar{x}' - \bar{x}_1 = 3,01 - 3,9013 = -0,8913$$

4.2.3.1.3. \bar{x}' 'nin varyans tahmini

σ_2^2 yanıtlayanlar tabakasının varyansı bilinmediği için yerine yanıtlayanlar tabakasından alınan altörnekten elde edilen varyans tahmini $s_2'^2 = 1,6865$ değeri kullanılacaktır. $\sigma^2 = \sigma_2^2$ olduğu için σ^2 'nin tahmini de $s^2 = s_2'^2 = 1,6865$ olur.

Örnek ortalamasının varyansı, $\frac{N}{N-1} = 1$ ve $\frac{N_2}{N_2-1} = 1$ olduğu varsayımı yapıldığı için $N=N-1$ ve $N_2=N_2-1$ dir ve $V(\bar{x}')$ formülünde σ_2^2 , σ^2 ve W_2 değerleri yerine tahminleri olan $s_2'^2 = 1,6865$, $s^2 = s_2'^2 = 1,6865$ ve $w_2=0,73$ değerleri kullanılacağından varyans formülü tahmine dönüşmüş olur.

$$\begin{aligned} v(\bar{x}') &= \frac{N-n}{N} \frac{s_2'^2}{n} + \frac{1}{n.N} (k-1)s_2'^2 \\ &= \frac{384-246}{384} \frac{1,6865}{246} + \frac{1}{246.384} (1,3-1)1,6865 \end{aligned}$$

$$v(\bar{x}') = 0,00247$$

Tek örnekleme yöntemiyle ortalamanın varyans tahmini 0,01655 olurken, Hansen ve Hurwitz'in çift örnekleme yöntemiyle ortalamasının varyans tahmini 0,00247'dir. olarak bulunmuştur. $v(\bar{x}') = 0,00247 < v(\bar{x}) = 0,01655$ olduğundan çift örnekleme yönteminin daha duyarlı yönteminin daha duyarlı bir yöntem olduğu söylenebilir.

4.2.3.2. Cochran'ın çift örnekleme yöntemi

Cochran'ın yönteminde ikinci soru grubu için de istenilen hesaplamalar birinci soru grubunda olduğu gibi iki varsayım altında yapılmıştır.

$$4.2.3.2.1. \frac{N}{N-1} = 1, \frac{N_2}{N_2-1} = 1 \text{ ve } S^2 = S_2^2 \text{ varsayımı altında Cochran'ın yöntemi}$$

Bu bölümde istenilen tüm hesaplamalar N ve N_2 'nin yeteri kadar geniş olduğu ve

$$\frac{N}{N-1} = 1, \frac{N_2}{N_2-1} = 1 \text{ 'e eşit olduğu varsayımı ile } S^2 = S_2^2 \text{ ele alınan yığının özellikleri}$$

ile yanıtlamayanların oluşturduğu yığının özelliklerinin aynı olduğu varsayımı altında yürütülmüştür.

4.2.3.2.1.1. k , n_2' ve n' 'in optimum değerlerinin hesaplanması

$n=246$ kişi üzerinden yürütülen anket çalışmasının yalnızca 70'inden yanıt alınabilmiştir. 70 kişiye ilişkin istatistik sonuçları aşağıdaki gibi verilmiştir.

Çizelge 4. 13. Posta anketini yanıtlayanlar ve yanıtlamayanlara ilişkin istatistik değerleri

Yanıtlayanlar	Yanıtlamayanlar
$n_1 = 70$	$n_2 = 176$
$w_1 = 0,28$	$w_2 = 0,72$
$\bar{x}_1 = 3,9013$	
$s_1^2 = 1,4168$	

$S^2 = S_2^2$ olduğu için optimum k aşağıdaki formülden hesaplanır.

$$k_{opt} = \sqrt{\frac{c_2 W_1}{c_0 + c_1 W_1}}$$

$$= \sqrt{\frac{5000000 \times 0,28}{400000 + 1500000 \times 0,28}}$$

$$k_{opt} = 1,3$$

Buradan yanıtlamayanlardan alınacak altörnek genişliği,

$$n_2' = n_2 / k_{OPT}$$

$$n_2' = 176 / 1,3 = 135$$

şeklinde bulunur.

Yanıtlamama olması durumunda yığından seçilecek ilk örnek genişliği n_{opt} ,

$$n_{opt} = \frac{N[1 + (k_{opt} - 1)W_2]}{N\left\{\left(\frac{N - n - 1}{N} \frac{1}{n}\right) + \frac{(k_{opt} - 1) W_2}{n}\right\} + 1}$$

formülünden bulunur. n_{opt} formülünde, optimum k değeri, W_2 yerine tahmini olan $w_2 = 0,73$ değeri, n ve N değerleri yerine yazılırsa,

$$n_{opt} = \frac{384[1 + (1,3 - 1)0,72]}{384\left\{\left(\frac{384 - 246}{384} \frac{1}{246}\right) + \frac{(1,3 - 1) 0,72}{246}\right\} + 1}$$

$$n_{opt} = 246$$

n_{opt} değeri hesaplanmış olur. Yanıtlamama olması durumunda yığından alınması gereken optimum örnek çapı 246 olarak bulunmuştur.

4.2.3.2.1.2. Yığın ortalamasının tahmini

Yanıtlamama olması durumu için çift örnekleme yöntemiyle yığın ortalamasının tahmini Hansen ve Hurwitz'in yönteminde olduğu gibi aşağıdaki formülle hesaplanır.

$$\bar{x}' = \frac{n_1}{n} \bar{x}_1 + \frac{n_2}{n} \bar{x}_2'$$

Posta anketini yanıtlayanlara ve yanıtlanmayan $n_2=176$ kişi arasından yüz-yüze görüşme yapılmak üzere seçilen $n_2'=135$ altörnek birimlerine ilişkin istatistik değerleri aşağıdaki gibidir.

Çizelge 4. 14. Posta anketini yanıtlayanlara ve yanıtlanmayanlar arasından seçilen altörnek birimlerine ilişkin istatistik değerleri

<i>Yanıtlayanlar</i>	<i>Yanıtlamayanlar</i>
$n_1 = 70$	$n_2 = 176$
$w_1 = 0,28$	$w_2 = 0,72$
$\bar{x}_1 = 3,9013$	$n_2' = 135$
$s_1^2 = 1,4168$	$\bar{x}_2' = 2,6686$
	$s_2^2 = 1,6865$

Buradan yığın ortalamasının yansız tahmini,

$$\bar{x}' = \frac{70}{262} 3,9013 + \frac{176}{262} 2,6686$$

$$\bar{x}' = 3,01$$

3,01 olarak hesaplanmış olur. Hansen ve Hurwitz'de olduğu gibi, $\frac{N}{N-1} = 1$,

$\frac{N_2}{N_2-1} = 1$ ve $S^2 = S_2^2$ varsayımı altında Cochran'ın çift örnekleme yöntemiyle

yığın ortalamasının tahminini ikinci soru grubu için 3,01 olup, “Kararsızım” şeklinde yorumlanır.

4.2.3.2.1.3. \bar{x}' 'nün varyans tahmini

Örnek ortalamasının varyansı aşağıdaki formülden hesaplanır.

$$V(\bar{x}') = \frac{N-n}{N} \frac{S^2}{n} + \frac{(k-1) W_2 S_2^2}{n}$$

$V(\bar{x}')$ formülünde, S_2^2 ve W_2 değerleri bilinmediğinden yerlerine tahminleri olan $s_2'^2 = 1,5815$, $w_2 = 0,73$ değeri ve bilinmeyen S^2 yığın varyansı için, $S^2 = S_2^2$ varsayımı altında, $s^2 = s_2'^2 = 1,6865$ tahmini kullanılır. Böylece $V(\bar{x}')$ formülü tahmine dönüşmüş olur;

$$\begin{aligned} v(\bar{x}') &= \frac{N-n}{N} \frac{s^2}{n} + \frac{(k-1) w_2 s_2'^2}{n} \\ &= \frac{384-246}{384} \frac{1,6865}{246} + \frac{(1,26-1) 0,72 (1,6865)}{246} \end{aligned}$$

$$v(\bar{x}') = 0,003747$$

Tek örnekleme yöntemiyle ortalamanın varyans tahmini 0,01655, Hansen ve Hurwitz'in çift örnekleme yöntemiyle ortalamanın varyans tahmini 0,00247 olarak bulunurken, Cochran'da bu değer 0,003747'dir.

Cochran (10) ve Rao (47) çıkarsamalarından bulunan varyans tahmini formülünü kullanarak yığın ortalamasının varyans tahmini,

$$\begin{aligned}
v(\bar{x}') &= \frac{(N-n)(n_1-1)}{N(n-1)} w_1 \frac{s_1^2}{n_1} + \frac{(N-1)(n_2-1) - (n-1)(n_2'-1)}{N(n-1)} w_2 \frac{s_2'^2}{n_2'} \\
&\quad + \frac{N-n}{N(n-1)} \left[w_1 (\bar{x}_1 - \bar{x}')^2 + w_2 (\bar{x}_2' - \bar{x}')^2 \right] \\
&= \frac{(384-246)(70-1)}{384(246-1)} 0,28 \frac{1,4168}{70} + \frac{(384-1)(176-1) - (246-1)(139-1)}{384(246-1)} 0,72 \frac{1,6865}{139} \\
&\quad + \frac{384-246}{384(246-1)} \left[0,28(3,9013 - 3,01)^2 + 0,72(2,6686 - 3,01)^2 \right]
\end{aligned}$$

= 0,00483 olarak elde edilir.

4.2.3.2.2. $\frac{N}{N-1} \neq 1$, $\frac{N_2}{N_2-1} \neq 1$ ve $S^2 \neq S_2^2$ varsayımı altında Cochran'ın yöntemi

Yığın ortalamasının yansız tahmini \bar{x}' 'nin bulunabilmesi için yanıtlamayanlar tabakasından alınacak altörnek çapının belirlenerek, altörnek birimleriyle yüz-yüze görüşme yapılarak veri toplanması gerekir.

4.2.3.2.2.1. k, n₂' ve n'in optimum değerlerinin hesaplanması

Araştırma yalnızca posta anketi yöntemine dayalı yürütülmüş (tek örnekleme yöntemi yapılmış) olsaydı yığın varyansı (S^2) tahmininin yanıtlayanların varyansına eşit olduğu ve S^2 yığın varyansının tahmini olan s^2 'nin $s_1^2=1,4168$ eşit olduğu kabul edilecekti.

Yanıtlamama olması durumunda, $S^2 \neq S_2^2$ varsayımı altında, n ve k'nın optimum değerlerini hesaplayabilmek için yanıtlayanlar ve yanıtlamayanlar tabakalarına ilişkin varyansların biliniyor olması istenir. Yanıtlayanlar tabakasının varyans tahmini $s_1^2=1,4168$ dir, ancak yanıtlamayanlar tabakasına ilişkin herhangi bir bilgi mevcut olmadığından varyans değeri (S_2^2) de bilinmemektedir.

S_2^2 'yi tahmin edebilmek amacıyla yanıtlamayanlar arasından eşit seçilme şansı vererek ve yerine koymadan seçim yöntemiyle tesadüfi 30 çaplı ön örnek seçilmesine karar verilmiştir. Öntest çalışması yanıtlamayan birimleriyle yüz-yüze görüşme yapılması şeklinde olmuş ve elde edilen bilgiler Çizelge 4.15.'te verilmiştir.

Çizelge 4. 15. Posta anketini yanıtlayanlar ve yanıtlamayanlar arasından seçilen $n_{20}=30$ ön örnek birimlerine ilişkin istatistik değerleri

Yanıtlayanlar	Yanıtlamayanlar
$n_1 = 70$	$n_2 = 176$
$w_1 = 0,28$	$w_2 = 0,72$
$\bar{x}_1 = 3,9013$	$n_{20} = 30$
$s_1^2 = 1,4168$	$\bar{x}_{20} = 2,4628$
	$s_{20}^2 = 0,7088$

Yığın varyansı S^2 'nin tahmini, posta anketini yanıtlayan $n=70$ kişinin verisine dayalı olarak $s_1^2=1,4168$ değerine eşit kabul edilmiştir. Yanıtlamayanlar tabakasının varyansı S_2^2 'nin tahmini, yanıtlamayanlardan seçilen öntest çalışması sonuçlarına dayalı olarak $s_{20}^2=0,7088$ olarak bulunmuştur. Buradan $\frac{N}{N-1} \neq 1$, $\frac{N_2}{N_2-1} \neq 1$ ve $S^2 \neq S_2^2$ varsayımı altında optimum n ve k hesaplanır.

Optimum k değeri,

$$k_{opt} = \sqrt{\frac{c_2(S^2 - W_2 S_2^2)}{S_2^2(c_0 + c_1 W_1)}}$$

$$k_{opt} = \sqrt{\frac{5000000(1,4168 - 0,72 \cdot 0,7088)}{0,7088(400000 + 1500000 \cdot 0,28)}}$$

$$k_{opt} = 2,79$$

2,79 olarak hesaplanır. Buradan yanıtlamayanlar tabakasından yüz-yüze görüşme yapılmak üzere seçilecek altörnek genişliği n_2'' 'ü,

$$n_2' = n_2 / k_{OPT}$$

$$n_2' = 176 / 2,79 = 63$$

şeklinde bulunur.

Yanıtlamama olması durumunda yığından seçilecek optimum ilk örnek genişliği formülünde $V = V(\bar{x}')$ değerinin belirlenmesi gerekir. $V(\bar{x}')$ formülünde, S^2 yerine $s_1^2 = 1,4168$ tahmin değeri, S_2^2 yerine $s_{2b}^2 = 0,7088$ tahmin değeri ve W_2 yerin $w_2 = 0,72$ tahmin değeri kullanıldığında varyans tahmin yapılmış olur. \bar{x}' 'nin varyans tahmini,

$$\begin{aligned} v(\bar{x}') &= \frac{N-n}{n} \frac{s^2}{n} + \frac{(k-1) w_2 s_2'^2}{n} \\ &= \frac{384-246}{384} \frac{1,4168}{246} + \frac{(2,79-1) 0,72 0,7088}{246} \\ &= 0,005783 \end{aligned}$$

şeklinde bulunur. Buradan istenilen n_{opt} değeri aşağıdaki gibi hesaplanır.

$$n_{opt} = \frac{384[1,4168 + (2,79-1) 0,72 0,7088]}{384 0,005783 + 1,4168}$$

$$n_{opt} = 246$$

Yanıtlamama olması durumunda yığından alınması gereken optimum ilk örnek genişliği, $\frac{N}{N-1}=1$ ve $\frac{N_2}{N_2-1}=1$ ile $S^2=S_2^2$ olduğu varsayımının yapıldığı durum ve yapılmadığı durum için aynı olup 246 olarak bulunmuştur.

4.2.3.2.2.2. Yığın ortalamasının tahmini

$\frac{N}{N-1} \neq 1$, $\frac{N_2}{N_2-1} \neq 1$ ve $S^2 \neq S_2^2$ varsayımı altında yanıtlamayanlar tabakasından

seçilecek altörnek genişliği $n_2' = 63$ olarak hesaplanmıştır. Yanıtlamayanlar arasından tesadüfi olarak seçilen 63 kişiyle yüz-yüze görüşme yapılmış ve bu kişilere ilişkin verilere dayalı olarak hesaplanan istatistik değerleri aşağıdaki gibi bulunmuştur.

Çizelge 4. 16. Altörnek birimlerine ilişkin istatistik değerleri

Ortalama (\bar{x}_2')	=	2,6336
Mod	=	2,00
Standart Sapma (s_2')	=	1,3158
Varyans ($s_2'^2$)	=	1,7314

$63 \times 17 = 1077$ veri üzerinden yanıtlamayanlar tabakasının ortalaması 2,6336 ve varyansı 1,7314 olarak tahmin edilmiştir.

Yığın ortalamasının çift örnekleme yöntemiyle tahmini Çizelge 4.17'de verilen değerlerin \bar{x}' formülünde yerine yazılmasıyla bulunur.

Çizelge 4. 17. Posta anketini yanıtlayanlar ve yanıtlamayanlar arasından seçilen altörnek birimlerine ilişkin istatistik değerleri

Yanıtlayanlar	Yanıtlamayanlar
$n_1 = 70$	$n_2 = 176$
$w_1 = 0,28$	$w_2 = 0,72$
$\bar{x}_1 = 3,9013$	$n_2' = 63$
$s_1^2 = 1,4168$	$\bar{x}_2' = 2,6336$
	$s_2'^2 = 1,7314$

Yanıtlamama olması durumu için çift örnekleme yöntemiyle yığın ortalamasının tahmini,

$$\bar{x}' = \frac{n_1}{n} \bar{x}_1 + \frac{n_2}{n} \bar{x}_2'$$

$$\bar{x}' = \frac{70}{246} 3,9013 + \frac{176}{246} 2,6336$$

$$\bar{x}' = 2,99$$

İkinci soru grubu için yığın ortalamasının tahmini 2,99 "Kararsızım" olarak bulunmuştur.

4.2.3.2.2. 3. \bar{x}' 'nün varyans tahmini

\bar{x}' 'nün varyansı ilk olarak,

$$V(\bar{x}') = \frac{N-n}{N} \frac{S^2}{n} + \frac{(k-1) W_2 S_2^2}{n}$$

formülden hesaplanabilir. $V(\bar{x}')$ varyans formülünde; S^2 , S_2^2 ve W_2 yığın parametreleri yerine sırasıyla s_1^2 , s_2^{02} ve w_2 tahmin değerleri kullanıldığında, varyans formülünün tahmine dönüştüğü daha önce belirtilmişti. $V(\bar{x}')$ 'in tahmini alt bölüm 4.2.2.1.'de 0,005783 olarak bulunmuştur.

Cochran (10) ve Rao (47) çıkarsamalarına dayalı olarak \bar{x}' 'nün varyans tahmini,

$$v(\bar{x}') =$$

$$\frac{(384-246)(70-1)}{384(246-1)} 0,28 \frac{1,4168}{70} + \frac{(384-1)(176-1) - (246-1)(63-1)}{384(246-1)} 0,72 \frac{1,7314}{63}$$

$$+ \frac{384 - 246}{384(246 - 1)} [0,28(3,9013 - 2,99)^2 + 0,72(2,6336 - 2,99)^2]$$

$$v(\bar{x}') = 0,001195$$

şeklinde bulunur.

4.2.3.3. Srinath'ın çift örnekleme yöntemi

Srinath'ın çift örnekleme yöntemiyle, ikinci soru grubu için \bar{X} yığın ortalaması tahmini ve tahminin varyansı hesaplanmaya çalışılacaktır.

Yanıtlamama olması durumunda, \bar{X} yığın ortalamasının yansız bir tahmininin yapılabilmesi için öncelikle, yanıtlamayanlar tabakasından ne kadar kişiyle yüz-yüze görüşme yapılacağına karar verilmesi gerekir.

4.2.3.3.1. k , n_2' ve n 'in optimum değerlerinin hesaplanması

Bu bölümde yanıtlamama olması durumunda yığından seçilmesi gereken örnek genişliği n ve altörnek genişliği n_2' hesaplanmaya çalışılacaktır. Bunun için öncelikle, Srinath'ın önermiş olduğu, k' altörnekleme kesrinin aşağıdaki formülden hesaplanacağı verilmiştir.

$$k'_{opt} = \sqrt{\frac{(S^2 - W_2 S_2^2) c_2 W_2^2}{S_2^2 (c_0 + c_1 W_1)}} - W_2$$

k'_{opt} formülü, $\frac{N}{N-1} \neq 1$, $\frac{N_2}{N_2-1} \neq 1$ ve $S^2 \neq S_2^2$ varsayımı altında Cochran'ın yöntemiyle hesaplanan k_{opt} 'a göre tanımlanacak olursa, aşağıdaki gibi yazılabilir.

$$k'_{opt} = (k_{opt} - 1) W_2$$

$k_{opt} = 2,79$ olarak bulunmuştur. Buradan, optimum k' değeri,

$$k'_{opt} = (k_{opt} - 1) W_2$$

$$k'_{opt} = (2,79-1) \cdot 0,73$$

$$k'_{opt} = 1,3$$

şeklinde hesaplanmış olur.

Yanıtlamayanlar tabakasından yüz-yüze görüşme yapılmak üzere seçilecek altörnek genişliği,

$$n_2' = n_2 / k'_{opt}$$

$$n_2' = 176 / 1,3 = 135$$

şeklinde elde edilir.

Yanıtlamama olması durumunda yığından alınması gereken optimum ilk örnek genişliği ise aşağıda verilen formülden hesaplanabilir.

$$n = \hat{n} \left(1 + \frac{k' S_2^2}{S^2} \right)$$

\hat{n} : Tam yanıt olması durumunda yığından alınacak örnek genişliği

$\hat{n}=246$ olduğu bilinmektedir. Burada, S_2^2 ve S^2 varyansları bilinmediği için, altbölüm 4.2.3.2.2.1.'de elde edilen hesaplamaların sonuçlarından, S^2 yerine $s_1^2=1,4168$ tahmin değeri, S_2^2 yerine $s_{20}^2= 0,7088$ tahmin değeri kullanılır. Yanıtlamama olması durumunda yığından seçilecek altörnek çapı,

$$n = \hat{n} \left(1 + \frac{k' S_2^2}{S^2} \right)$$

$$n = 246 \left(1 + \frac{1,3 \cdot 0,7088}{1,4168} \right)$$

$$n = 406$$

406 olarak bulunmuştur. Bulunan bu sonuç, yığındaki tüm birimlerle görüşme yapılması gerektiğini belirtmektedir.

4.2.3.3.2. Yığın ortalamasının tahmini

Yanıtlamayanlar örneğinden yüz-yüze görüşme yapılmak üzere seçilecek altörnek birimlerinin sayısı $n_2'=135$ olarak hesaplanmıştır. İkinci soru grubu için bulunan bu değer, Hansen ve Hurwitz'in yönteminde bulunan altörnek genişliğiyle aynı olduğu için, ortalama tahmini de aynı bulunmuştur. Dolayısıyla yığın ortalamasının yansız tahmini,

$$\bar{x}' = \frac{70}{262} 3,9013 + \frac{176}{262} 2,6686$$

$$\bar{x}' = 3,01$$

3,01 olarak hesaplanmış olur.

4.2.3.3.3. \bar{x}' 'nün varyans tahmini

Yukarıda hesaplanan k' , S_2^2 ve S^2 'nin tahminleri, $V(\bar{x}')$ eşitliğiyle tanımlanmış varyans formülünde yerine yazılacak olduğunda,

$$V(\bar{x}') = \left(\frac{N-n}{Nn}\right)S^2 + \frac{k'S_2^2}{n}$$

$$V(\bar{x}') = \left(\frac{384-246}{384 \cdot 246}\right)1,4168 + \frac{1,3 \cdot 0,7088}{246}$$

$$V(\bar{x}') = 0,0058$$

varyans formülüne de ulaşmış olacaktır.

İkinci soru grubuna ilişkin tek örnekleme ve çift örnekleme yöntemleriyle elde edilen sonuç istatistikleri Çizelge 4.18.'de özetlenmiştir.

Çizelge 4. 18. İkinci soru grubu sonuç istatistikleri

YÖNTEMLER	Varsayımlar	k_{opt}	n_2'	n	Yığın Ortalamasının tahmini	Yığın Ortalamasının tahmininin Varyans tahmini	Cochran (10) ve Rao (46) çıkarsamalarından Ortalamanın Varyans tahmini
Tek örnekleme Yöntemi	-	-	-	246	3,9	0,016550	-
Hansen ve Hurwitz'in Çift Örnekleme Yöntemi	$\frac{N}{N-1} = 1, \frac{N_2}{N_2-1} = 1$ ve $\sigma^2 = \sigma_2^2$ varsayımı	1,3	135	299	3,01	0,002470	-
Cochran'ın Çift Örnekleme Yöntemi	$\frac{N}{N-1} = 1, \frac{N_2}{N_2-1} = 1$ ve $S^2 = S_2^2$ varsayımı	1,3	135	246	3,01	0,003747	0,00483
	$\frac{N}{N-1} \neq 1, \frac{N_2}{N_2-1} \neq 1$ ve $S^2 \neq S_2^2$ varsayımı	2,79	63	246	2,99	0,005783	0,001195
Srinath'ın Çift Örnekleme Yöntemi	-	1,3	135	406	3,01	0,00580	-

Çizelgeye bakıldığında, yığın ortalamasının tahmini tek örnekleme yöntemiyle 3,9 olurken, Hansen ve Hurwitz'in, $\frac{N}{N-1}=1$, $\frac{N_2}{N_2-1}=1$ ve $S^2=S_2^2$ varsayımı altında

Cochran'ın ve Srinath'ın çift örnekleme yöntemleriyle 3,01, $\frac{N}{N-1} \neq 1$, $\frac{N_2}{N_2-1} \neq 1$

ve $S^2 \neq S_2^2$ varsayımı altında ise 2,99 olarak bulunmuştur. Buradan, ikinci soru grubuyla ölçülmek istenen çalışanların örgüte bağlılığı durumu, tek örnekleme yöntemine dayalı olarak, çalışanların örgüte bağlı olduğunu göstermiştir. Çift örnekleme yöntemiyle ise, örgüte bağlılık konusunda çalışanların kararsız olduğu tespit edilmiştir.

Çizelge 4.18.'de ortalama tahminin varyans tahminlerine bakıldığında, en yüksek varyans değerinin tek örnekleme yöntemine ait olduğu görülmektedir. Dolayısıyla en düşük duyarlılığa sahip örnekleme yöntemi tek örnekleme yöntemidir. Buradan, çift örnekleme yöntemlerinin tek örnekleme yönteminden daha duyarlı olduğu söylenebilir. Birinci soru grubunda olduğu gibi burada da, örnekleme yöntemleri arasında en düşük varyanslı ve en duyarlı örnekleme yöntemi $\frac{N}{N-1}=1$, $\frac{N_2}{N_2-1}=1$

ve $\sigma^2 = \sigma_2^2$ varsayımıyla Hansen ve Hurwitz'in çift örnekleme yöntemi olmuştur. Burada da yine, birinci soru grubunda gözleendiği gibi, tüm çift örnekleme yöntemleriyle bulunan varyans tahmini değerleri arasında çok da büyük farklılıklar olmadığı görülmektedir.

İkinci soru grubu için belirlenen n_2' altörnek çaplarına bakıldığında, en düşük altörnek çapının Cochran'ın $\frac{N}{N-1} \neq 1$, $\frac{N_2}{N_2-1} \neq 1$ ve $S^2 \neq S_2^2$ varsayımı altında

hesaplanan altörnek çapına ait olduğu görülmektedir. Bu yöntemle $n_2'=63$ olarak bulunmuş ve toplam görüşme yapılan kişi sayısı 133 kişi olmuştur. Diğer çift örnekleme yöntemlerinde ise altörnek çapı 135, toplam görüşme yapılan kişi sayısı ise 205 kişi olmuştur. Buradan, çift örnekleme yöntemleri arasında Cochran'ın

$\frac{N}{N-1} \neq 1$, $\frac{N_2}{N_2-1} \neq 1$ ve $S^2 \neq S_2^2$ varsayımı altında kullanılan yöntemin, birinci

soru grubunda olduğu gibi, daha düşük maliyetli olduğu görülmüştür.



6. SONUÇ VE ÖNERİLER

Yanıtlamamaya sayım ve anketlerde sıklıkla rastlanılmaktadır. Bu nedenle anket çalışmasının çok iyi düzenlenmesi ve yürütülmesi gerekir. Ancak en iyi düzenlenmiş anketlerde dahi yanıtlamama ortaya çıkmaktadır, soruna çözüm getirebilmek için özel tahmin yöntemlerine ihtiyaç vardır (28). Yanıtlamama istatistiksel analizlerde ortaya çıkan önemli problemlerden birisidir. Yanıtlamamanın büyük oranda oluşu yapılan çalışmanın güvenilirliğini azaltmaktadır.

Düşük yanıt oranı bulguların geçerliliğine ve genelleştirilmesine zarar vereceğinden bir sorun olarak kabul edilen yanıtlamamanın giderilmesi ya da minimum yapılması için çeşitli çalışmaların yürütülmesi gereklidir. Yanıtlamama düzeyinin çok düşük olduğu durumlarda bile yanıtlamama hataları kategorik olarak ele alınmalı ve yanıtlamayanlardan veri toplanılmaya çalışılmalıdır (40).

Bazen örnek araştırmaları için yanıt oranı %15'ten az ise o araştırmanın kullanılamaz olduğu söylenir. Ancak çalışma kapsamında hala yanıtlanmayan birimlerden bilgi alınabilme imkanı söz konusu olabiliyorsa, yüksek yanıtlamama oranı çok büyük tehlike arz etmez (40).

Colombo (33), yanıt oranının pazarlama ve reklamcılık araştırmalarında azalmakta olduğunu, yanıtlamama yanının ise artış gösterdiğini belirtmiştir. Genellikle posta anketleri için yanıt oranının %20, telefon anketleri için ise %50'lerin altında olduğu, elektronik ortamda yürütülen araştırmaların ve gönderilen e-mail anket çalışmalarının yaygın hale geldiği vurgulanmıştır.

Tez'de, yanıtlamama olması durumunda çift örnekleme yönteminin ampirik bir çalışması yapılmış, çalışmamızın sonuçları tezin uygulama bölümünde verilmiştir. Nevşehir Tekel Fabrikası çalışanlarına iş ortamı profilinin tespit edilmesi amacıyla posta anketi uygulaması yapılmıştır. Tekel Fabrikasında yürütülen posta anketi çalışmasında yanıtlamama durumu söz konusu olmuştur, bu durum, Hansen ve Hurwitz (9), Cochran (10) ve Srinath (11)'in yanıtlamama için çift örnekleme

yönteminin bir uygulamasının yapılabilmesine imkan vermiştir. Ancak, Nevşehir Tekel Fabrikası çalışanlarına ilişkin yanıt oranlarının tahmini yapılamadığı için, Deming'in yönteminin uygulaması gerçekleştirilememiştir.

Yürütülen çalışmada, yanıtlamayan birimlerinden tesadüfi seçilen altörnek birimlerinin tümünden ilk aşamada yanıt alınabildiği için, ikiden fazla aşama altörnek alınması yöntemleri uygulanmasına gerek kalmamıştır.

Fabrika çalışanlarına yöneltilen soruları içeren anket formu demografik sorular, birinci soru grubu ve ikinci soru grubundan oluşmaktadır. Birinci soru grubu işletme çalışanları üzerinde örgütsel adaletin uygulanıp uygulanmadığını ölçerken, ikinci soru grubu işletme çalışanlarının kendilerini örgüte bağlı hissedip hissetmediklerini ölçmek için tertiplenmiştir. Tezin uygulama bölümü, birinci soru grubu ve ikinci soru grubu üzerinden yürütülen iki ayrı çalışmanın sonuçlarını içermiştir. Araştırma konusu iki değişken (örgütsel adaletin uygulanıp uygulanmadığı, çalışanların kendilerini örgüte bağlı hissedip hissetmedikleri değişkenleri) olduğu için, yığından seçilecek örnek çapı iki değişkene dayalı olarak ayrı ayrı hesaplanmış, yığına ilişkin tahminler de iki ayrı örnek üzerinden yürütülmüştür.

Çalışanlarının adreslerine anket formları postalanmış ancak anketleri yanıtlayanların oranı birinci soru grubu için %30'dan daha düşük olmuştur. Fabrikaya gidilerek postalanan anket formlarının neden yanıtlanmadığı araştırıldığında, Nevşehir Tekel Fabrikasının özelleştirme aşamasında olması ve fabrikada çalışan sözleşmeli işçilerin işten çıkarılma kaygısı taşımaları nedeniyle, anket uygulamasının fabrika yönetimi tarafından yapıldığını düşündükleri, anket formunda yer alan soruları çok özel buldukları ve yanıtlama durumunda işten çıkarılma tehlikesi taşıdıkları öğrenilmiştir. Anketin amacı, kimler tarafından yürütülmüş olduğu anket formunun en üstünde belirtilmiş olduğu halde, yürütülen çalışmaya karşı bir güvensizlik söz konusu olmuştur.

Anketörlerin işçilerle yaptıkları yüz-yüze görüşmelerde, anketin amacı, kimler tarafından yürütüldüğü, ad soyad bilgilerinin istenilmediği, sonuçların ne şekilde değerlendirileceği açıklanarak yürütülen çalışmaya olan güven sağlanmış, çalışma sonuçlarının kendilerine ulaştırılacağı belirtilmiştir. Böylece altörnek birimleriyle yapılan yüz-yüze görüşmelerde tam yanıt alınmıştır. Yüz-yüze görüşme yapılmamış olsaydı, belki de işçiler posta anketlerinde soruları çok özel bulmaları nedeniyle, gerçek bilgilerini yansıtmayacak yanıtlar vereceklerdi, bu da yanıt hatalarına yol açacaktı.

Buradan, posta anketlerini yanıtlayanların genelinin memurlar, yanıtlamayanların ise sözleşmeli işçiler olduğu tespit edilmiştir. Yanıtlamayanların yanıtlayanlardan araştırma değişkenleri bakımından farklı özellik taşıdıkları gözlenmiştir. Birinci soru grubu için yanıtlamama oranı %73 iken, ikinci soru grubu için %72'dir. Yanıt oranının çok düşük olması nedeniyle, yanıtlamama yanın ise giderilmesi gerektiği için yanıtlamayan birimlerinden altörnek alınacağına karar verilmiştir.

Uygulama çalışmasında, birinci ve ikinci soru grubu için öncelikli olarak, yalnızca yanıtlayanlar verisine dayalı tek örnekleme yöntemi sonuç istatistiklerine yer verilmiştir. İkinci aşama olarak ise; Hansen ve Hurwitz'in yöntemi $\frac{N}{N-1}=1$,

$$\frac{N_2}{N_2-1}=1 \text{ ve } \sigma^2 = \sigma_2^2 \text{ varsayımı altında, Cochran'ın yöntemi } \frac{N}{N-1}=1, \frac{N_2}{N_2-1}=1$$

ve $S^2=S_2^2$ ile $\frac{N}{N-1} \neq 1$, $\frac{N_2}{N_2-1} \neq 1$ ve $S^2 \neq S_2^2$ varsayımı altında ayrı ayrı ve

Srinath'ın çift örnekleme yöntemlerinin sonuç istatistiklerine yer verilmiştir.

Birinci soru grubu için yığın ortalamasının tahmini tek örnekleme yöntemiyle 3,696

olurken, $\frac{N}{N-1}=1$, $\frac{N_2}{N_2-1}=1$ ve $\sigma^2 = \sigma_2^2$ varsayımı altında Hansen ve

Hurwitz'in ve $\frac{N}{N-1}=1$, $\frac{N_2}{N_2-1}=1$ ve $S^2=S_2^2$ varsayımı altında Cochran'ın çift

örnekleme yöntemleriyle 2,75 olarak, $\frac{N}{N-1} \neq 1$, $\frac{N_2}{N_2-1} \neq 1$ ve $S^2 \neq S_2^2$ varsayımı altında Cochran'da ise 2,86 olarak bulunmuştur. Srinath'ın çift örnekleme yöntemiyle $k_{opt}=0,59$ bulunduğu için, bu yöntem kullanılamamıştır. Birinci soru grubu çalışanlar üzerinde işyerinde örgütsel adaletin uygulanıp uygulanmadığını ölçen sorulardan oluşmaktadır. Dolayısıyla, tek örnekleme yöntemine dayalı olarak, çalışanların örgütsel adaletin uygulandığı görüşünde olduğu, çift örnekleme yöntemiyle ise, örgütsel adaletin uygulandığı konusunda çalışanların kararsız oldukları tespit edilmiştir.

İkinci soru grubu için yığın ortalamasının tahmini tek örnekleme yöntemiyle 3,9 olurken, $\frac{N}{N-1}=1$, $\frac{N_2}{N_2-1}=1$ ve $\sigma^2 = \sigma_2^2$ varsayımı altında Hansen ve Hurwitz'in, $\frac{N}{N-1}=1$, $\frac{N_2}{N_2-1}=1$ ve $S^2=S_2^2$ varsayımı altında Cochran'ın ve Srinath'ın çift örnekleme yöntemleriyle 3,01, $\frac{N}{N-1} \neq 1$, $\frac{N_2}{N_2-1} \neq 1$ ve $S^2 \neq S_2^2$ varsayımı altında ise 2,99 olmuştur. Buradan, tek örnekleme yöntemine dayalı olarak çalışanların örgüte bağlı olduğu, çift örnekleme yöntemine dayalı olarak ise çalışanların örgüte bağlılık konusunda kararsız olduğu yorumu yapılır.

Yalnızca yanıtlayanlar verisine dayalı yapılacak tahminlerin genellikle yanlı olacağı belirtilmişti. Yürütülen uygulama çalışmasıyla da örnek birimlerinin yalnızca %27 veya %28'lik bölümünden elde edilen verilere dayalı yapılan tahminlerin yanlı olduğu ve bu tahminlerin yığının özelliklerini yansıtmadığı birinci ve ikinci soru grubunun her ikisi için de söylenebilir.

Birinci soru grubu ve ikinci soru grubu için, tek ve çift örnekleme yöntemleri sonucunda hesaplanan ortalama tahminin varyans tahminlerine bakıldığında, her iki soru grubu için de tek örnekleme yönteminde bulunan varyans değerinin çift örnekleme yöntemleriyle bulunan varyans değerlerinden yüksek olduğu görülmüştür.

Dolayısıyla, yanıtlamama olması durumunda, çift örnekleme yönteminin tek örnekleme yönteminden daha duyarlı olduğu söylenebilir.

Uygulama çalışmasında örnekleme yöntemleri arasında en düşük varyanslı olanı, her iki soru grubu için de, $\frac{N}{N-1}=1$, $\frac{N_2}{N_2-1}=1$ ve $\sigma^2 = \sigma_2^2$ varsayımıyla Hansen ve

Hurwitz'in çift örnekleme yöntemi olduğu için, en duyarlı örnekleme yöntemi Hansen ve Hurwitz'in çift örnekleme yöntemidir denir. Ancak, tüm çift örnekleme yöntemleriyle bulunan varyans tahmini değerleri arasında çok da büyük farklılıklar olmadığı da görülmüştür.

Çift örnekleme yöntemleriyle hesaplanan \bar{x} 'nin varyans tahmini değerleri arasında çok büyük farklılıklar olmadığı gözlenirken, altörnek çapında belirgin farklılık olduğu görülmüştür. Altörnek çapı, birinci soru grubu için; $\frac{N}{N-1}=1$, $\frac{N_2}{N_2-1}=1$ ve

$S^2=S_2^2$ varsayımı altında Cochran ile $\frac{N}{N-1}=1$, $\frac{N_2}{N_2-1}=1$ ve $\sigma^2 = \sigma_2^2$ varsayımı

altında Hansen ve Hurwitz'in yöntemiyle 148, $\frac{N}{N-1} \neq 1$, $\frac{N_2}{N_2-1} \neq 1$ ve $S^2 \neq S_2^2$

varsayımıyla Cochran'ın yöntemiyle 107 bulunmuştur. İkinci soru grubu için;

$\frac{N}{N-1}=1$, $\frac{N_2}{N_2-1}=1$ ve $\sigma^2 = \sigma_2^2$ varsayımı altında Hansen ve Hurwitz ile

$\frac{N}{N-1}=1$, $\frac{N_2}{N_2-1}=1$ ve $S^2=S_2^2$ varsayımı altında Cochran'ın ve Srinath'ın

yöntemiyle 135, $\frac{N}{N-1} \neq 1$, $\frac{N_2}{N_2-1} \neq 1$ ve $S^2 \neq S_2^2$ varsayımıyla Cochran'ın

yöntemiyle 63 bulunmuştur.

Altörnek çapı, çift örnekleme yöntemleri arasında Cochran'ın $\frac{N}{N-1} \neq 1$, $\frac{N_2}{N_2-1} \neq 1$

ve $S^2 \neq S_2^2$ varsayımı altında kullanılan çift örnekleme yöntemiyle, her iki soru grubu

için de, diğer yöntemlere kıyasla daha düşük bulunmuştur. Cochran'ın $\frac{N}{N-1} \neq 1$, $\frac{N_2}{N_2-1} \neq 1$ ve $S^2 \neq S_2^2$ varsayımı altında kullanılan yöntemi, diğer çift örnekleme yöntemlerine kıyasla, az sayıda yanıtlamayan birimle görüşme yapılmasını önermiş ve daha düşük maliyet gerektirmiştir.

Yürütülen uygulama çalışması analiz sonuçlarına dayalı olarak, aynı doğrulukta sonuçlar veren ve düşük maliyet sağlayan Cochran'ın $\frac{N}{N-1} \neq 1$, $\frac{N_2}{N_2-1} \neq 1$ ve $S^2 \neq S_2^2$ varsayımı altında kullanılan yöntemin gelecekte yürütülecek çalışmalarda tercih nedeni olabileceği söylenebilir. Ayrıca örnekleme yöntemleri arasında en duyarlı olanı, her iki soru grubu için de, $\frac{N}{N-1} = 1$, $\frac{N_2}{N_2-1} = 1$ ve $\sigma^2 = \sigma_2^2$ varsayımıyla Hansen ve Hurwitz'in çift örnekleme yöntemi olduğu için de, bu yöntemin araştırmanın duyarlılığı açısından tercih edilebileceği söylenebilir.

Elde edilen sonuçlara göre; yanıtlamama olması durumunun yana neden olduğu, yanıtlamama yanının ise giderilmesi gerektiği, bunun için önerilen çift örnekleme yönteminin kullanılabilirliği görülmüştür.

KAYNAKLAR

1. Groves, R. M., "Survey Errors And Survey Costs", *John Wiley & Sons Inc.*, New York (1989).
2. Rubin, D. B., "Multiple imputation for nonresponse in surveys", *John Wiley & Sons Inc.*, New York (1987).
3. Mack, S., Peteroni, R., "Overview of SIPP Nonresponse Research", *Fifth International Workshop On Household Survey Non-Response*, September 26-28, Ottawa, Ontario (1994).
4. Ayhan, H. Ö., "Survey Nonresponse Models And Applications In Turkey", *Conference On Methodological Issues In Official Statistics*, Stockholm, Sweden Announcement (October 1998).
5. Leigh, J. H., Martin, C. R., " "Don't Know" Item Nonresponse In A Telephone Survey: Effects Of Question Form And Respondent Charecteristics", *Journal Of Marketing Research*, 24: 418-424 (November 1987).
6. Kish, L., "Survey Sampling", *John Wiley & Sons Inc.*, New York (1965).
7. Moser, C. A., Kalton, G., "Survey Methods In Social Investigation", *Heinemann Educational Books*, London (1971).
8. Ayhan, H. Ö., "Türkiye Doğurganlık Araştırmasında (1978) Yanıtlamama Kaynakları ve Yanlılığı", *H. Ü. Nüfus Etütleri Enstitüsü Nüfus Bilim Dergisi*, 5 (1983).
9. Hansen, M.H., Hurwitz, W. N., "The Problem Of Nonresponse In Sample Surveys", *Journal Of The American Statistical Association*, 41: 517-529 (1946).
10. Cochran, W. G., "Sampling Techniques", *John Wiley & Sons Inc.*, New York (1977).
11. Srinath, K. P., "Multiphase Sampling in Nonresponse Problems", *Journal of the American Statistical Association*, 66: 583-586 (1971).

12. Deming, W. E., "On A Probability Mechanism To Attain An Economic Balance Between The Resultant Error Of Response And The Bias Of Nonresponse", *Journal of the American Statistical Association*, 48: 743-772 (1953).
13. Çıngı, H., "Örnekleme Kuramı", *H.Ü. Fen Fakültesi Yayınları*, Ankara (1994)
14. Koç, İ., "Soru Kağıdı Hazırlama ve Soru Sorma Teknikleri Ders Programı", *D.İ.E. Anketör Eğitim Merkezi Temel Eğitim Programı*, Ankara (2002).
15. Ulusoy, M., "Türkiye Nüfus Ve Nüfus Sorunları Araştırması Evli Kadın Anketinin Araştırma Metodolojisi Açısından Değerlendirilmesi", Yüksek Lisans Tezi, *H. Ü. Nüfus Etütleri Enstitüsü*, Ankara (1973)
16. Balcı, A., "Sosyal Bilimlerde Araştırma", *Pegem Yayınevi*, Ankara (2001)
17. Şenesen Ü., "İşletme ve İktisat İçin İstatistik", *Literatür Yayınları*, İstanbul (2000).
18. Darwin, C.C., "Determinants of Item Nonresponse", *Survey Research Techniques Seminar Fall 2002*, www.personal.psu.edu/faculty/d/r/drj10/Cruz.pdf (25.04.2003).
19. Groves, R. M., Biemer, P. P., Lyberg, L. E., Massey, J. T., Nicholis W. L. and Waksberg, J., "Telephone Survey Methodology", *John Wiley & Sons Inc.*, New York, 191-212 (1987).
20. Esin, A., Bakır, M. A., Aydın, C. ve Gürbüzsel, E., "Temel Örnekleme Yöntemleri", *Literatür Yayınları*, İstanbul (2001).
21. Serper, Ö., "Uygulamalı İstatistik", *Ezgi Kitabevi*, Bursa (2000).
22. Schuman, H., Presser, S., "Questions And Answers In Attitude Surveys: Experiments On Question Form, Wording And Context", *Academic Press*, New York (1995).
23. Dedeş, H., "Anketör Eğitimi, Alan Çalışması ve Pre-Test", Türk Hava Kurumu Personeli İçin Saha Araştırma Metot ve Teknikleriyle Ölçme ve Değerlendirme Eğitim Programı, *D.İ.E. Eğitim Merkezi Koordinatörlüğü Yayını*, Ankara (2002).

24. Francis, J. D., Busch, L., "What We Now Know About "I Don't Knows"", *The Public Opinion Quarterly*, 39 (2): 207-218 (1975).
25. Faulkenberry, G.D., Mason, R., "Characteristics of Nonopinion And No Opinion Response Groups", *The Public Opinion Quarterly*, 42 (4): 533-543 (1978).
26. Atrostic, B. K., Bates N., Burt G. and Silberstein, A., "Nonresponse In U. S. Government House Hold Surveys: Consistent Measures, Recent Trends And New Insights", *Journal Of Official Statistics*, 17, No. 2, 209-226 (2001).
27. Lin, B., Jones, C. A., "Some Issue In Conducting Customer Satisfaction Survey", *Journal Of Marketing Practice: Applied Marketing Science*, 3 (1): 4-13 (1997).
28. Särndal, C. E., Swensson, B. and Wretman, J., "Model Assisted Survey Sampling", *Springer- Verlag*, New York (1992).
29. Little, R., Rubin, D. B., "Statistical Analysis With Missing Data", *John Wiley & Sons Inc.*, New York (1987).
30. Sherman, R. P., "Test Of Certain Types Of Ignorable Nonresponse In Surveys Subject To Item Nonresponse Or Attration", *American Journal Of Political Science*, 44 (2): 362-388 (April 2002).
31. <http://www.Bls.Census.Gov/Sipp/Qp/C5.Htm>, 02.01.2002
32. Churchill, G.A., "Marketing Research", *Dryden Press*, Florida (1996).
33. Colombo, R., "A Model For Diagnosing And Reducing Nonresponse Bias", *Journal Of Advertising Research*, 40 (1/2): 85-99 (2000).
34. Levy, P. S., Lemeshow, S., "Sampling Of Populations: Methods And Applications, Solutions Manual", *John Wiley & Sons Inc.*, New York (1991).
35. http://wwwlib.umi.com/dissertations/preview_page/NQ47726/13, 22.08.2001
36. Baş, T., "Anket", *Seçkin Yayınları*, Ankara (2001).

37. Çıngı, H., "Araştırma Yöntemleri", *H.Ü. Fen Fakültesi İstatistik Bölümü Ders Notları*, Ankara (1994).
38. Thompson, S. K., "Sampling", *John Wiley & Sons Inc.*, New York (1992).
39. Raj, D., "Sampling Theory", *McGraw-Hill*, New York (1971).
40. Teksoy, N., "Cevaplamama Hatalarının ve İkamelerin Tahminler Üzerindeki Etkileri", Uzmanlık Tezi, *T.C. Başbakanlık D.İ.E. Yayını*, Ankara (1991).
41. Madow, W. G. and Singh, B., "Incomplete Data In Sample Surveys: The Panel On Incomplete Data In Sample Surveys", *Academic Press*, New York, 93-95 (1983).
42. Rao, P. S. R. S., "Incomplete Data In Sample Surveys: The Panel On Incomplete Data In Sample Surveys", *Academic Press*, New York, 97-105 (1983).
43. El-Badry, M. A., "A Sampling Procedure For Mailed Questionnaires", *Journal Of The American Statistical Association*, 51: 209-227 (1956).
44. Sukhatme, P.V., Sukhatme, B.V., Sukhatme, S. And Asok, C., 1984, "Sampling Theory Of Surveys", *Indian Society Of Agricultural Statistics And Iowa State University Press.*, Ames (1984).
45. Politz, A. and Simmons, W.R., "An Attempt To Get The "Not At Homes" Into The Sample Without Call-Backs", *Journal of the American Statistical Association*, 44: 9-31 (1949).
46. Birnbaum, Z.W. and Sirken, M.G. , "Bias Due To Non-Availability In Sampling Surveys", *Journal Of The American Statistical Association*, 45: 98-111 (1950).
47. Rao, J. N. K., "On Double Sampling For Stratification And Analytical Surveys", *Biometrika* 60: 12-133 (1973).
48. Betlehem, J.G. and Kersten, H.M.P., "On The Treatment Of Nonresponse In Sample Survey", *Journal Of Official Statistics*, 1: 287-300 (1985).



EKLER

EK1: Anket Formu



EK-1: Anket Formu**İŞLETMELERDE PERSONELE SAĞLANAN DESTEK VE BAĞLILIK
ARAŞTIRMASI**

Değerli çalışanlar,

Yüksek lisans tezimin uygulama kısmını oluşturan bu anket, fabrika personeline sağlanan destek ve bağlılık düzeyini belirlemeye yönelik bir araştırmadır. Araştırmanın değeri ve başarısı tümüyle sizin katılımınıza bağlıdır.

Araştırmada anket formunu dolduranın kimliği sorulmamaktadır. Bu nedenle, anket formu üzerine kimliğinizi yazmayınız. Anket formunu doldurduktan sonra, size anket formu ile birlikte gönderilen üzerinde adresimiz yazılı zarfa koyarak postalayınız.

Katkılarınız için şimdiden teşekkür ederiz.

Tez yöneticisi:
Prof. Dr. A. Alptekin ESİN
Gazi Üniversitesi Fen Fakültesi
İstatistik Bölüm Başkanı

Tezi uygulayan:
Aylin ALKAYA
Gazi Üniversitesi Fen Fakültesi
İstatistik Bölüm Yüksek lisans
öğrencisi

ANKET SORULARI

Cinsiyetiniz	() Erkek () Bayan
Medeni Durumunuz	() Evli () Bekar () Boşanmış
Eğitim Düzeyiniz	() Okuma-yazma bilmiyor () Herhangi bir okul bitirmemiş () İlkokul () Ortaokul () Lise () Üniversite () Lisansüstü
Yaşınız
İşletmedeki çalışma süreniz (ay veya yıl olarak belirtiniz) (ay) (yıl)

1. Aşağıdaki ifadelere katılma derecenizi yanındaki boşluğa X işareti ile belirleyiniz.

Sıra	İfade	Kesinlikle aynı fikirde değilim	Aynı fikirde değilim	Kararsızım	Aynı fikirdeyim	Kesinlikle aynı fikirdeyim
1	Bu işletmede kendimi rahat ve huzurlu hissediyorum.					
2	Bu işletmede kendimi önemli hissedirim.					
3	Benimle ilgili bir karar alınmadan önce bunun beni nasıl etkileyeceği dikkate alınmaz.					
4	İşletmede başarılı olduğum zaman yöneticiler benimle övünürler.					
5	En zor işi bile yapsam bu hiç kimsenin umurunda değildir.					
6	Fırsat buldukları anda benim aleyhimde kararlar alabilirler.					
7	İşletme, benim sıkıntılı zamanlarımda bana yardımcı olur.					
8	Şikayetlerim yönetim tarafından hiç dikkate alınmaz.					
9	Bu işletmede benim fikirlerime değer verilir.					
10	Benim yerime daha ucuz ücretle çalışacak birini bulsalar hemen işe alırlar.					
11	Çalışma koşullarımın daha iyi olması için yöneticiler bana yardımcı olur.					
12	Her an işime son verilme tehlikesi hissediyorum.					
13	Bu işletmede benimle kimse ilgilenmez.					
14	Yönetim, benim görev sadakatimi bilir ve görevimi iyi yaptığımı düşünür.					
15	Benim yararına olacağını bilsem bile bu işletmeden ayrılmayı düşünmem.					
16	Bu işletmeye devam etme sorumluluğu hissetmiyorum.					

17	Bu işletmeden ayrılırsam kendimi suçlu hissederim.					
18	Şu an işletmedeki sorumluluklarım ve görevlerim nedeniyle ayrılmam yanlış olur.					
19	Bu işletmeye çok şey borçluyum.					

2. Aşağıdaki ifadelere katılma derecenizi yanındaki boşluğa X işareti ile belirleyiniz.

Sıra	İfade	Kesinlikle aynı fikirde değilim	Aynı fikirde değilim	Kararsızım	Aynı fikirdeyim	Kesinlikle aynı fikirdeyim
1	Burada, kendimi ailenin bir ferdi gibi hissetmiyorum.					
2	Kendimi bu işletmeye duygusal olarak bağlı hissediyorum.					
3	Bu işletmede çalışmanın benim için çok özel bir anlamı var.					
4	Kendimi bu işletmeye ait olarak hissediyorum.					
5	Benim sadakatim bu işletmede hiç dikkate alınmaz.					
6	Bu işletmede çalıştığımı arkadaşlarıma ve aileme gururla söylüyorum.					
7	Emekli olana kadar bu işletmede çalışmayı çok isterim.					
8	Bu işletmenin sıkıntılarını kendi sıkıntılarım gibi kabul ediyorum.					
9	İşletmemle ilgili olarak dışarıda konuşmaktan büyük bir zevk duyuyorum.					
10	Eğer bir gün kendi isteğimle ayrılmak isteseydim, bu durumda bile bu işletmeden ayrılmam çok zor olurdu.					
11	Bu işletmeden ayrılırsam hayatımdaki pek çok şey daha kötü olur.					
12	Bu işletmede uzun vadeli olarak çalışmayı düşünmüyorum.					
13	Bu işletmeden ayrılmak şu an için benim açımdan zor.					
14	Bu işletmede kalmamın tek nedeni bu işletmenin bana sağladığı imkanları dışarıda bulamayacak olmamdır.					
15	Dışarıdan bana gelecek aynı pozisyondaki iş teklifleri beni ilgilendirmiyor.					

3. Aşağıdaki ifadelere katılma derecenizi yanındaki boşluğa X işareti ile belirleyiniz.

Sıra	İfade	Kesinlikle aynı fikirde değilim	Aynı fikirde değilim	Kararsızım	Aynı fikirdeyim	Kesinlikle aynı fikirdeyim
1	İşletmeye gelirken her gün pozitif (olumlu) düşüncelerle geliyorum.					
2	İşletmede genel olarak negatif (olumsuz-çatışmacı) bir çalışma ortamı bulunmaktadır.					

4. İşletmedeki konumunuz	<input type="checkbox"/> Memur <input type="checkbox"/> İşçi
İşletmedeki göreviniz	<input type="checkbox"/> Orta düzey yönetici (Bölüm müdürleri / yardımcıları vb.) <input type="checkbox"/> Alt düzey yönetici (Ustabaşı-tekniker vb.) <input type="checkbox"/> Sade bir çalışan

Araştırmaya katıldığınız için teşekkür ederiz...



EK-2. <http://www.Bls.Census.Gov/Sipp/Qp/C5.Htm> İnternet Kaynađı

5. NONRESPONSE ERROR

Three kinds of nonresponse occur in SIPP: household (unit) nonresponse, person (within unit) nonresponse, and item nonresponse. Survey procedures are designed to minimize all three kinds of nonresponse, and weighting and imputation procedures have been developed to try to minimize errors resulting from differences between units and persons that do respond and those that do not. Section 5.1 describes the steps taken in SIPP to maximize response rates. Sections 5.2 through 5.4 present data on levels and trends for each kind of nonresponse and on the characteristics of respondents and nonrespondents. Nonresponse indicators for SIPP and other surveys are compared. Finally, Section 5.5 describes the results of experiments with procedures designed to improve response rates in SIPP. The weighting and imputation procedures designed to minimize the effects of nonresponse are covered in Chapter 8.

5.1 Steps to maximize response rates

Wave 1 response

Several steps are taken to encourage response to the first wave interview:

- An advance letter from the Director of the Census Bureau explains the authority for and purposes of the survey and urges participation.
- Field representatives (FRs) carry cards and portfolios identifying them as Census Bureau employees.
- If no one is home at the time of the first visit, interviewers try to determine the best time for a callback, either by asking neighbors or by telephoning.
- Supervising field representatives and supervisors may be called in to convert refusals.
- More flexibility has been introduced into the interviewing schedule for later panels to allow more time to follow up on refusals.
- FRs are provided with talking points to explain SIPP and how it is used in policy decisions. The talking points target different segments of the population including: the elderly, the poor, the wealthy, and the middle-income groups.

Wave 2 and beyond

Keeping response rates high in longitudinal surveys is especially difficult, because units and persons who respond initially can be lost in later waves for various reasons. Steps taken in SIPP to minimize attrition after the initial wave include:

- A letter from the Census Director is sent to responding households in advance of each wave thanking them for their support and urging continued participation.
- Respondents are given a brochure containing some interesting results from the survey.

- Interviewers are instructed to call respondents and schedule interviews at times that are convenient for them.
- To assist in locating households that move, contact persons who would be likely to know the new locations are identified at the time of the first interview. Interviewers are instructed on how to trace movers by contacting new occupants of the housing unit, the Post Office, and other likely sources of information.

Success in completing interviews is an important factor in evaluating interviewer performance. The standards that have been established for household response rates, calculated on an annual basis, are shown in Section 4.4. The Census Bureau has continued to review these standards to ensure that they truly reflect the FRs' efforts in completing interviews with both movers and nonmovers.

5.2 Household nonresponse

Household or unit nonresponse occurs at the initial wave when all members of an eligible household refuse to participate in the survey or when no one can be contacted after repeated attempts. Except in the 1985 panel, there have been no attempts to contact wave 1 noninterview households in subsequent waves. When this was tried in the second wave of the 1985 panel, relatively few households were brought back into the sample and the processing of these households turned out to be difficult and resulted in delays in releasing data products, so the practice was discontinued (Jean and McArthur, 1987).

After wave 1, sampled households comprise those households that contain original sample persons. Household nonresponse from these households can occur for the same reasons as in wave 1. In addition, these households may be temporarily unavailable because a sample person has moved, and the household cannot be located, or because the household is located more than 100 miles from a SIPP sample PSU and cannot be interviewed by telephone. Whenever possible, a household that is not interviewed in the second or a subsequent wave is visited once more in the next wave to attempt an interview. If that attempt fails, there are no further attempts to interview the household in later waves.

Nonresponse rates

The calculation of response (or nonresponse) rates is complicated by the fact that the exact number of eligible households after the first wave is not known. Households interviewed in wave 1 may split up to form additional eligible households, or may leave the survey population entirely. When all members of an interviewed household move and cannot be located, they may account for zero, one, or two or more eligible households in subsequent waves. If all members of the household leave the survey population, there will be no eligible households. If they split up and move to different housing units, there can be two or more households.

The accuracy of reported response rates can also be affected by interviewer misclassification of noninterviews. If a noninterview household is incorrectly classified as not eligible for interview, it will be incorrectly excluded from both the numerator and denominator of the calculated nonresponse rate, and vice versa. Reinterview results for the years 1984 to 1987 show that only 3.5 percent of noninterviews were incorrectly classified, so the effect of such errors on reported response rates is minimal (Pennie, 1988).

Type A and Type D nonresponse

Table 5.1 shows household noninterview rates by panel and wave for the 1984-1993 SIPP panels for Type A and Type D noninterviews. Type A noninterviews occur when interviews are not obtained for known eligible households. Type D noninterviews occur when some or all members of a household leave and cannot be traced or move more than 100 miles from a SIPP PSU and cannot be interviewed by telephone. The rates shown are cumulative. The sample loss rates after wave 1 include an adjustment for unobserved growth resulting from splits of households that were Type A noninterviews in the first wave and later waves. Details of the calculation of the rates shown in Table 5.1 are given by Nelson et al. (1987). Total sample loss is highest at the first and second interviews, and most of the early losses are due to Type A nonresponse. The number of additional Type A noninterviews is quite small after the first four waves. Most households that have participated up to that point stick it out to the end. Losses due to moves (Type D noninterviews) continue at a fairly constant rate throughout the duration of each panel.

Rates from 1984-1993

Table 5.1 shows a consistent pattern in the wave nonresponse rates for SIPP panels (see Citro and Kalton, 1993, pp. 100-102). Between 1984 and 1990, the rate of loss is about 5-8 percent of eligible households at wave 1, an additional 4-6 percent of eligible households at wave 2, another 2-3 percent in each of waves 3-5, and less than 1 percent in each subsequent wave thereafter. For these panels, by wave 6 (after 2 years of interviewing), cumulative sample loss from SIPP is 18-20 percent of eligible households; by wave 8 it is 21-22 percent (Citro and Kalton, 1993). The 1992 and 1993 panels experienced higher household nonresponse. The loss rates were around 9 percent at wave 1, 22 percent at wave 6, and 25 percent at wave 8.

The causes for the increase in household nonresponse in recent panels is unknown. One suggestion from the FRs is that maximum telephone interviewing has an adverse effect on Type A noninterview rates. Investigations by King (1995) and James (1994) found no evidence to support this suggestion. However, Type D nonresponse rate was found to be high when telephone interviewing was used heavily.

It is relevant to note here that several other major demographic household surveys conducted by the Census Bureau also experienced an increase in Type A noninterview rates and refusal rates in the period 1990-1995 (Davey, 1996). The surveys investigated include: the CPS, the National Crime and Victimization Survey, National Health Interview Survey (NHIS), Consumer Expenditure (CE) Quarterly, and Consumer Expenditure (CE) Diary. Comparisons of the annual average Type A noninterview rate between 1990 and 1995 showed increases ranging from 1.4 percent for the NHIS to almost 13 percent for the CE Diary. The increases in the annual average refusal rates ranged from 1.4 percent to 4.1 percent.

Rates relative to other surveys

Because of important differences in survey content and design features, it is difficult and may possibly be misleading to compare the SIPP response experience with that of other longitudinal surveys, let alone cross-sectional surveys. Nelson et al. (1987) made a careful comparison of sample losses in SIPP, the Income Survey Development Program (ISDP), the National Longitudinal Surveys (NLS), the National Medical Care Utilization and Expenditure Survey, and the Panel Study of Income Dynamics (PSID). They concluded that: "In summary, SIPP sample losses appear to be comparable to those experienced by other longitudinal surveys. We see the same pattern repeated in all these surveys; that is, a heavy initial loss which increases at a much slower rate throughout the subsequent interviews until it levels off." (Nelson et al., 1987, p.636). Further information on nonresponse and other nonsampling errors in longitudinal surveys can be found in Kasprzyk et al. (1989).

Table 5.1 Household noninterview¹ and sample loss² rates: 1984-1993 panels³

Wave	1984 Panel ³			1985 Panel			1986 Panel			1987 Panel			1988 Panel		
	Type A	Type D	Loss	Type A	Type D	Loss	Type A	Type D	Loss	Type A	Type D	Loss	Type A	Type D	Loss
1	4.9	–	4.9	6.7	–	6.7	7.3	–	7.3	6.7	–	6.7	7.5	–	7.5
2	8.3	1.0	9.4	8.5	2.1	10.8	10.8	1.5	13.4	11.1	1.5	12.6	11.4	1.5	13.1
3	10.2	1.9	12.3	10.2	2.7	13.2	12.6	2.3	15.2	11.5	2.6	14.2	12.0	2.3	14.7
4	12.1	2.9	15.4	12.4	3.4	16.3	13.8	3.0	17.1	12.3	3.3	15.9	13.0	3.0	16.5
5	13.4	3.5	17.4	14.0	4.1	18.8	15.2	3.7	19.3	13.7	4.1	18.1	13.9	3.3	17.8
6	14.9	4.1	19.4	14.2	4.8	19.7	15.2	4.3	20.0	13.6	4.9	18.9	13.6	4.0	18.3
7	15.6	4.9	21.0	14.4	5.2	20.5	15.3	4.8	20.7	13.6	4.9	19.0			
8	15.8	5.7	22.0	14.4	5.5	20.8									
9	15.8	5.7	22.3												

Wave	1989 Panel			1990 Panel			1991 Panel			1992 Panel			1993 Panel		
	Type A	Type D	Loss	Type A	Type D	Loss	Type A	Type D	Loss	Type A	Type D	Loss	Type A	Type D	Loss
1	7.6	–	7.6	7.3	–	7.3	8.4	–	8.4	9.3	–	9.3	8.9	–	8.9
2	11.0	1.4	12.5	10.9	1.5	12.6	12.3	1.5	13.9	12.8	1.7	14.6	12.3	1.7	14.2
3	11.2	2.3	13.8	11.5	2.6	14.4	13.1	2.7	16.1	13.1	2.8	16.4	12.9	2.8	16.0
4				12.5	3.4	16.5	13.6	3.6	17.7	13.8	3.6	18.0	13.9	3.5	17.9
5				13.6	4.6	18.8	14.5	4.2	19.3	14.9	4.7	20.3	14.9	4.4	19.9
6				14.1	5.3	20.2	14.4	5.1	20.3	15.3	5.4	21.6	15.9	5.5	22.2
7				14.3	5.9	21.1	14.7	5.6	21.0	16.0	5.9	23.0	17.1	6.2	24.0
8				14.4	5.9	21.3	14.5	5.9	21.4	16.9	6.7	24.7	17.5	6.9	25.1
9										17.7	7.3	26.2	18.1	7.5	26.5
10										17.5	7.6	26.7			

¹Type A noninterviews consist of households occupied by persons eligible for interview and for whom a questionnaire would have been completed if an interview had been obtained. Reasons for Type A noninterview include: no one at home in spite of repeated visits, temporarily absent during the entire interview period, refusal, and unable to locate a sample unit. Type D noninterviews consist of households of original sample persons who are living at an unknown new address or at an address located more than 100 miles from a SIPP PSU, provided a telephone interview is not conducted.

²The sample loss rate consists of cumulative noninterview rates adjusted for unobserved growth in the Type A noninterview units (created by splits).

³Differences in rates for the 1984 panel compared to subsequent panels may be due in part to differences in the panel designs.

Source: Data for the 1984-1988 panels came from Kasprzyk (1988) and Bowie (1990).

Refusal

Most Type A noninterviews are the result of refusals. Refusals accounted for between 70 and 76 percent of wave noninterviews in the first four panels (Bowie, 1988b, p.8). Some data on reasons for refusals in the first six waves of the 1984 panel are given by Nelson et al. (1987). Only one reason for refusing was coded for each household even though multiple reasons may have been given. For wave 1, the most frequently cited reasons were:

Reason given	Percent of households that refused
Not interested in participating	18.7
No time, too busy	14.7
Invasion of privacy	9.9
Voluntary survey	9.3
Offended by income questions, too personal	9.1
Didn't believe information was confidential	3.4

Of those who refused in the second wave, 6.2 percent said they had not understood that the interviewer would be returning for another interview. In the third and later waves, about one-fourth of the refusers said that they did not want to answer any more questions.

Characteristics of nonresponding households

For obvious reasons, relatively little information is available about the characteristics of sample households that are not interviewed in wave 1 of a panel. It is possible, however, to compare the characteristics of households that are interviewed in the first wave and drop out later with the characteristics of households that continue to participate. The Census Bureau's SIPP Nonresponse Workgroup (1988) compared distributions based on first-wave characteristics of household reference persons for sample households interviewed in all of the first five waves of the 1984 panel and sample households interviewed in the first wave but not in the fifth wave (households in the second group may or may not have been interviewed in waves 2 to 4). Types of households more likely to be noninterviews in the fifth wave included those that were: occupied by renters, located in Standard Metropolitan Statistical Areas (SMSAs) with a population of 500,000 or more, or had reference persons in the age group 15 to 24. No statistically significant differences were observed for distributions of households by household monthly income, sex of the reference person, or receipt/nonreceipt of cash or noncash benefits.

Mack and Waite (1995) proposed a nonresponse follow-up study to obtain more information about wave 1 noninterviews. The plan was to mail a short questionnaire to all wave 1 nonrespondents to obtain basic information about household membership, income, and program participation. The objective is to use the information to refine the noninterview adjustments to the sampling weight. This proposal was implemented for the 1996 panel and results are forthcoming.

5.3 Person nonresponse

Person nonresponse occurs when a member of an interviewed household refuses to participate or is unavailable for interview and a proxy interview is not obtained. Original sample persons can miss any number of waves and still reenter the sample as long as their households have not missed two or more consecutive interviews in the interim. Most analyses of person nonresponse subsequent to wave 1 include all persons not interviewed in each wave, whether or not other members of their household were interviewed in that wave.

As discussed at length in Section 8, the presence of person nonresponse in interviewed households makes it difficult to decide what combinations of imputation and estimation procedures are best suited for undertaking various kinds of longitudinal analyses with SIPP data.

Nonresponse pattern

Table 5.2 (from Lepkowski et al., 1989) shows the person response patterns for the eight interviews in the 1984 panel. Panel members are original sample persons who responded for all eight interviews for which they were eligible (i.e., for which they were in the SIPP universe). Nonpanel members are original sample persons who failed to respond to one or more interviews when they were eligible. Total nonrespondents are those who did not respond to any of the eight interviews. Table 5.2 shows that 70.4 percent of the SIPP sample were interviewed in all interviews in which they were eligible. Persons who were interviewed for one or more consecutive interviews and then dropped out for all of the remaining interviews, known as attrition cases, accounted for 14 percent. Persons with other response patterns, who might be called "in and outers," accounted for 7.2 percent.

Table 5.3 shows the nonresponse patterns for panel and nonpanel members from waves 1 through 7 of the 1985, 1986, 1987, and the 1991 panels. This table does not include an estimate for total nonrespondents. Therefore, the percentages are not directly comparable to Table 5.2. Table 5.3 shows that about 76 to 80 percent of the SIPP sample who responded in wave 1 continued to respond to all seven waves in which they are eligible. Persons who were interviewed for wave 1, or more consecutive interviews, and then attrite from all remaining interviews accounted for about 13-16 percent. Persons with other nonattrition patterns accounted for about 6-7 percent. About half of the persons who were Type Z nonrespondents at wave 1 failed to respond at any of the seven waves. Type Z nonrespondents are members of otherwise cooperating households for whom no information is obtained, either in person or by proxy.

Movers

Not surprisingly, person nonresponse is higher for persons who move than for those who are residents of the same housing unit for the duration of the panel. Table 5.4 compares person nonresponse for movers and nonmovers in waves 1 to 5 of the 1984 panel, for rotation groups 1, 2, and 3. About one-fifth of the original sample persons moved at least once between waves 1 and 5. One or more interviews were missed for about one-third of the movers, as compared with less than one-fifth of the nonmovers (Jean and McArthur, 1987). For the 1991 panel, Mack et al., 1995 also found that at wave 8, the nonresponse rate for movers was 22.5 percent, a rate that doubled the rate of 10.3 percent for nonmovers.

Table 5.2 Person response (1) or nonresponse (0) patterns for the eight interviews in the 1984 SIPP panel research file

Person response pattern	Percent
Panel members	70.4
Responded to all 8 interviews	67.0
Responded to all Interviews prior to death	1.5
Responded to all Interviews except when:	
in an institution	0.7
in the Armed Forces	0.4
abroad	0.6
Other panel members	0.2
Nonpanel members	24.8
Attrition nonresponse patterns:	14.0
11111110	1.4
11111100	1.2
11111000	1.5
11110000	2.0
11100000	2.2
11000000	2.6
10000000	3.1
In and Outers:	7.2
Other patterns with only one missing interview	4.3
Other patterns with two or more missing interviews	2.9
Nonpanel members with missing interviews who were known to have died, become institutionalized, entered the Armed Forces, or moved abroad	0.3
Nonpanel members with missing interviews who left for other reasons	3.3
Total nonrespondents¹	4.9
Total	100.0
Total number of sample persons	35,027

¹Estimated -- the number of total nonrespondents is not known; all that is available is the number of total nonresponding households, from which the number of persons has been crudely estimated by multiplying the average number of sample persons in responding households by the number of nonresponding households.

Source: Lepkowski et al. (1989).

Table 5.3 Nonresponse patterns for panel and nonpanel members in waves 1-7 of the 1985-87 and 1991 panels

Pattern group	1985	1986	1987	1991
	Panel	Panel	Panel	Panel
	Percent	Percent	Percent	Percent
Panel members	76.1	77.5	79.5	77.0
A. Responded on all 7 waves	73.7	74.7	76.9	73.8
B. Ineligible on some waves:	2.4	2.8	2.6	3.2
Died	1.0	1.1	1.2	1.4
Entered institution	0.5	0.6	0.5	0.8
Entered Armed Forces barracks	0.3	0.3	0.2	0.2
Moved abroad	0.6	0.7	0.7	0.8
Reason ineligible not listed	0.0	0.0	0.0	0.0
Nonpanel members	23.9	22.5	20.5	23.0
C. Eligibility for all 7 waves	23.6	22.3	20.3	22.7
1. Attrition nonresponse only at :	15.8	14.4	12.6	13.3
Wave 2	5.1	5.0	4.7	4.6
Wave 3	3.4	2.9	2.3	2.5
Wave 4	2.8	2.1	1.9	1.8
Wave 5	1.8	2.0	1.6	1.4
Wave 6	1.3	1.1	1.1	1.2
Wave 7	1.4	1.3	1.1	1.8
2. Nonattrition nonresponse only:	5.7	5.7	5.6	6.8
1 wave	4.5	4.6	4.3	5.2
2 waves	0.9	0.9	1.0	1.1
3-5 waves	0.3	0.2	0.3	0.5
3. Nonattrition and attrition nonresponse:	1.3	1.3	1.2	1.4
At least one wave of each	1.3	1.3	1.2	1.4
4. Type Z noninterview at wave 1:	0.8	0.8	0.9	1.2
Noninterview all waves	0.4	0.4	0.4	0.6
Interviewed at least one wave	0.4	0.4	0.4	0.6
D. Some ineligibility	0.3	0.3	0.2	0.3
Total	100.0	100.0	100.0	100.0
Number of sample members	30,605	23,157	30,769	28,184

Table 5.4 Response status for movers and nonmovers in waves 1 to 5: 1984 panel¹

Response status	Total		Movers		Nonmovers	
	Number	Percent	Number	Percent	Number	Percent
Five interviews	19,878	79.1	3,485	68.8	16,393	81.7
Missing at least interview 5	4,222	16.8	1,148	22.6	3,074	15.3
All other	1,038	4.1	436	8.6	602	3.0
Total	25,138	100.0	5,069	100.0	20,069	100.0

¹Persons 15+ in rotation Groups 1, 2, and 3.

Source: Jean and McArthur (1987).

The higher nonresponse rates among movers is a concern because the SIPP sample shows relatively high mobility. Waite (1995a) reported that over 25 percent of households with three or more persons would have at least one person moving out of their households in a 4-month period. On average, 41 percent of the new noninterview cases reported in each wave of the 1992 panel are unlocated movers. Furthermore, in wave 8 of the 1991 panel, the noninterview rate for movers in poverty was twice the rate for persons not in poverty. In contrast, nonresponse due to refusals was higher for persons not in poverty than for persons in poverty.

Characteristics of nonrespondents

A study by McArthur (1987), based on all 9 waves of the 1984 panel, compared the characteristics of persons for whom all interviews were completed with the characteristics of persons for whom there were one or more noninterviews. Table 5.5 shows selected results from this study, and corresponding data from the waves 1-7 of the 1992 panel. The characteristics of persons with higher than average person nonresponse are similar in both panels. They are: residence in rented housing units, members of racial minorities, children and other relatives of the reference person, aged 15 to 24, never married, and had no savings account or other assets. Higher nonresponse rate are also common among residence in large SMSAs.

5.4 Item nonresponse

Item nonresponse occurs when a respondent refuses or is unable to answer a question. It can also occur when a FR fails to ask a question that should have been asked or asks the question but fails to record the answer. Missing data for an item can also be the result of a data entry clerk failing to key the answer to a question, or an answer being deleted in editing because it is inconsistent with answers to other items.

Table 5.5 Selected characteristics of original sample persons by their interview experience for the full 1984 panel and for the first 7 waves of the 1992 panel

Characteristics of persons in wave 1	1984				1992			
	Interview experience (% distribution)				Interview experience (% distribution)			
	Completed all	Missed last 2	Other ¹		Completed all	Missed last 2	Other ¹	
Missed 1			Missed 2+	Missed 1			Missed 2+	
Relationship:								
Ref. persons	37.3	33.1	33.9	30.7	36.3	31.3	30.5	31.2
Primary ind.	12.7	14.7	13.4	12.0	13.9	18.0	15.4	17.9
Spouse	31.3	23.9	25.4	23.9	29.4	21.3	22.4	21.0
Child	13.9	17.9	18.4	24.1	14.2	14.4	17.8	14.8
All other	4.8	10.3	9.0	9.4	6.3	15.0	13.9	15.1
Age:								
15-24	18.2	24.2	25.3	29.5	16.3	23.4	25.5	23.8
25-34	22.1	21.4	22.6	22.9	21.1	23.3	25.3	23.4
35-44	17.3	14.4	17.0	14.9	21.1	16.2	16.4	16.1
45-64	27.0	23.2	21.5	23.2	26.0	19.9	18.0	19.8
65+	15.4	16.8	13.6	9.5	15.4	17.2	14.9	16.9
Race:								
White	88.5	83.8	82.9	78.2	87.0	80.9	78.9	80.8
Black	9.2	12.9	13.9	17.6	9.6	14.5	15.6	14.6
Other	2.3	3.3	3.2	4.2	3.5	4.6	5.6	4.6
Living quarters:								
Owned	72.7	62.9	65.6	59.8	71.9	55.6	57.8	55.2
Rented	24.8	35.1	32.1	37.3	25.9	41.6	39.1	42.1
Rent free	2.4	2.0	2.3	2.8	2.3	2.8	3.1	2.8
Marital status:								
Never married	20.8	28.8	28.3	33.6	23.4	33.0	32.8	33.5
Married	63.5	52.2	53.3	50.4	59.9	46.9	47.9	46.6
Other	15.8	18.9	18.5	15.9	16.7	20.0	19.3	19.9
Savings account:								
Yes	59.9	51.1	51.9	51.0	55.0	42.8	43.3	42.5
No	40.1	48.9	48.1	49.0	45.0	57.1	56.7	57.5

¹Interview experience categories are mutually exclusive. At least one of the last two interviews was completed for persons in the other category.

Source: Data for the 1984 panel was adapted from McArthur (1987).

Item nonresponse can be analyzed either on a cross-sectional or a longitudinal basis. Cross-sectional item nonresponse refers to missing data items for persons interviewed in a single wave. Longitudinal item refers to data items common to two or more waves of interviewing: there is said to be longitudinal item nonresponse for such an item if the person was interviewed in all the relevant waves and data are missing for one or more of those waves. This section examines the extent of item nonresponse in SIPP and how it compares with other surveys. The methods used to compensate for cross-sectional and longitudinal item nonresponse are discussed in Chapter 8.

Core items

Table 5.6 shows the item nonresponse rates for selected core items in three categories: labor force activity; income reciprocity or asset ownership; and income amounts for five panels between 1984 and 1993. The rates for the first two categories are quite low and do not show much change over the panels. Nonresponse rates for income amounts, especially self-employment income, and interest and dividends, are substantially higher, indicating that respondents who report income reciprocity or asset ownership are not always able or willing to report amounts. For interest, as shown by the rates in parentheses for the 1984-1986 panels, the majority of respondents who did not report amounts of interest were willing to report the balances in their interest-bearing accounts, thus providing a reasonable basis for imputation of the amounts of interest. For dividends, separate questions are asked for dividends paid directly to shareholders and those credited to reinvestment or margin accounts (shown in parenthesis). Nonresponse rates are considerably higher for the latter category.

Table 5.7 shows longitudinal nonresponse rates for selected income amounts based on 12 months and 32 months of data from the 1984 panels. For income types appearing in both Tables 5.6 and 5.7 the overall nonresponse rates are lower in Table 5.6. For example, Food Stamps had an average cross-sectional nonresponse rate of 3.6 percent in the 1984 panel as compared with a longitudinal nonresponse rate of 8.1 percent for the first 12 months of that panel, and 7.2 percent for all 32 months of the panel. The 32-month estimate appears lower because it included only panel members who completed all 8 waves of the survey. Table 5.7 also shows that only 1.9 percent of Food Stamp recipients had no Food Stamps amounts reported for any of the 12 months, and 3.0 percent had no amounts reported for any of 32 months. Thus, for many Food Stamp recipients with some missing values for Food Stamp benefit amounts, it would be possible to impute these values based on data for other months.

Items on asset amounts

The highest item nonresponse rates experienced in SIPP have been for questions on asset amounts. Table 5.8 shows nonresponse rates for selected asset amounts collected in waves 3 and 4 of the 1984 SIPP panel, wave 4 of the 1986 panel and, for comparative purposes, from the Income Survey Development Program (ISDP). SIPP nonresponse rates for these items range from 13 to 42 percent, lower than the corresponding rates in the ISDP, but still much higher than rates for most other types of questions. Curtin, et al., 1989 have compared item nonresponse for questions on asset ownership and amounts in SIPP (data for waves 3 and 4 of the 1984 panel), the Panel Study of Income Dynamics (1984 Wealth Supplement), and the 1983 Survey of Consumer Finances (SCF). Nonresponse on the ownership questions is low for all types of assets in all three surveys. For the asset categories included in all three surveys, item nonresponse rates for questions on amounts were, without exception, higher for SIPP than for either of the other two surveys.

Table 5.6 Nonresponse rates for selected SIPP core items by panel

Question	1984	1985	1986	1992 ³	1993 ³
Labor force activity:					
Identification of weeks absent without pay (item 4)	0.1	(Z)	0.1	0.3	0.2
Identification of weeks with a job or business (item 6a)	2.2	2.0	2.5	0.2	0.1
Presence of weeks looking or on layoff (item 7a)	1.0	1.3	2.0	0.1	0.1
Identification of weeks looking or on layoff (item 7b)	3.2	2.4	2.9	0.2	0.1
Income reciprocity or asset ownership:					
Social Security	0.6	0.6	1.0	0.1	0.1
Unemployment compensation	0.1	0.1	0.2	0.2	0.1
Food Stamps	0.3	0.4	0.5	0.6	0.6
Savings accounts	1.0	0.9	0.9	0.8	1.0
Shares of stock	1.3	1.4	1.5	1.3	1.5
Income amounts:					
Hourly wage rate	9.5	10.4	10.8	7.2	7.7
Monthly wage and salary	6.2	7.2	8.4	4.0	4.2
Self-employment salary or draw	14.0	16.9	14.6	12.4	13.5
Social Security	8.8	9.5	10.0	14.0	14.7
Unemployment compensation	9.1	9.7	9.9	9.2	10.7
Food Stamps	3.6	4.1	4.4	6.4	5.4
Interest ¹	34.6(24.2)	29.8(28.9)	30.8(30.2)	10.3	10.2
Dividends ²	9.4(30.7)	10.5(30.5)	9.4(29.1)	7.5	7.6

Z = Less than .05 percent.

¹Figure in parentheses is the nonresponse rate on balance in the account. This question was asked on the 34.6 percent that did not provide an estimate of the amount of interest received.

²The figure in parentheses is the nonresponse rate for dividends credited to accounts.

³The rates shown for labor force activity items in 1992 and 1993 refer to wave 1 only.

Source: Rates for panel 1984-86 adapted from Bowie (1986).

Table 5.7 Longitudinal item nonresponse rates for amounts of selected income types: 1984 SIPP panel 12-month summary¹ and 32-month average²

Income type	All amounts reported	One or more amounts not reported	One or more but not all amounts not reported	No amounts reported
12 months				
Hourly wage rate	83.0	17.0	9.0	8.0
Social Security	82.8	17.2	13.1	4.1
Private pension	78.8	21.8	13.6	8.2
AFDC	91.0	9.0	5.6	3.4
Food Stamps	91.9	8.1	6.2	1.9
Unemployment compensation	87.9	12.1	4.0	8.0
Federal SSI	88.0	12.0	7.6	4.4
32 months				
Social Security	87.7	12.4	8.6	3.8
AFDC	92.1	7.9	4.8	3.1
Food Stamps	92.7	7.2	4.2	3.0
Unemployment compensation	86.6	13.4	4.5	8.9
Federal SSI	90.2	9.8	4.3	5.5

¹These rates are based on the total number of persons with reciprocity in one or more of the 12 months. Also these rates do not reflect imputations made to type Z person noninterviews.

²These rates are rates of missing data based on panel members included in the 1984 SIPP longitudinal file. The rates include imputations due to item nonresponse only. Type Z imputations are not included. Data is adapted from Pennell (1993), Table 4.3).

Source: Data for the 12-month summary was adapted from Kasprzyk and Herriot (1986) and data for the 32-month summary was adapted from Pennell (1993), Table 4-3.

Table 5.8 Item nonresponse rates for asset amounts in SIPP and the ISDP

Asset type	Percent nonresponse		
	SIPP		ISDP
	1984 waves 3 and 4	1986 wave 4	
Amount in savings accounts	16.8	22.8	24.9
Amount in checking accounts	13.3	21.2	23.1
Amount in bonds and government securities	25.9	23.1	32.2
Market value of stocks and mutual fund shares	41.5	36.9	65.8
Debt on stocks and mutual fund shares	41.1	38.6	87.3
Face value of U.S. Savings Bonds	24.9	24.9	35.8
Value of rental property	33.5	31.3	39.9
Value on own business	37.9	41.8	55.3
Debt on own business	28.8	31.5	50.4

Sources: Data from ISDP and SIPP 1984 Panel from U.S. Bureau of the Census (1986c).
Data from SIPP 1986 Panel from unpublished tabulations.

Topical modules

Turning to item nonresponse for the topical modules, we note first that there are some interviewed persons who do not provide any data, either directly or through a proxy respondent, for a topical module for which they are eligible. As shown in Table 5.9 for the 1984-1993 panels, complete nonresponse to topical modules ranges from about 3.0 to 9.0 percent. There is a tendency for this to increase moderately over the life of a panel. Topics covered in later waves of a panel are doubly disadvantaged: a smaller proportion of households and persons are interviewed, and of those who are interviewed, a slightly higher proportion are nonrespondents to the topical modules. By wave 8 of the 1993 panel, the topical module nonresponse rate for the school enrollment and financing module is 9.3 percent.

Item nonresponse rate in SIPP compares favorably to CPS

A question of considerable interest to users of SIPP income data is how the quality of those data compare with the quality of income data from the annual Current Population Survey (CPS) March Income Supplement. This is not an easy question to answer. A full analysis requires consideration of all sources of error that affect SIPP and CPS estimates, as well as differences in coverage and definitions between the two surveys.

Table 5.9 Person nonresponse rates for topical modules

Panel	Wave	Percent of persons with no topical module data	Type of module
1984	4	3.3	Fixed: Assets and liabilities Variable: Retirement and pensions, housing, energy
	5	3.1	Variable: Child care and support, welfare history, working, reservation wage
	6	4.2	Fixed: Annual income, taxes, employee benefits, education Variable: Training (ETA)
	8	4.5	Fixed: Marital, fertility and migration history Variable: Household relationships, work expense, support of persons not in household
	9	5.0	Fixed: Annual income, taxes, IRAs, education
1985	3	2.8	Fixed: Assets and liabilities
1990	2	5.0	Fixed: Education and work training history
	5	6.6	Fixed: Education financing
1991	2	5.0	Fixed: Education and training history
	5	6.0	Fixed: Education financing
	6	4.0	Variable: Extended measure of well-being
	8	6.5	Fixed: School enrollment and financing
1992	2	4.0	Fixed: Education and work
1993	8	9.3	Fixed: School enrollment and financing

Source: Bowie (1986, 1988a).

Looking only at item nonresponse rates for income amounts, SIPP has a clear advantage, as shown in Tables 5.10 and 5.11. Moreover, the SIPP sample persons who do not respond to particular income items are not necessarily the same ones in each quarter (see Table 5.7), so in many instances of item nonresponse, part-year data will be available for use in imputing missing values. In the March 1985 CPS, 5.2 percent of eligible households were not interviewed. In the interviewed households, data for the income supplement either were not obtained or could not be tabulated because of matching failures for 9.4 percent of the eligible persons. Combining these two rates indicates that the item nonresponse rates shown in Tables 5.10 and 5.11 for CPS apply to an interviewed population already subject to a 14.1 percent overall person nonresponse rate.

Table 5.10 Item nonresponse rates¹ for income amounts in the 1984 SIPP panel and March 1985 CPS, for selected income types

Income type	SIPP 1st quarter 1984 monthly average	SIPP 2nd quarter 1984 monthly average	SIPP 3rd quarter 1984 monthly average	SIPP 4th quarter 1984 monthly average	CPS March 1985
Wage and salary	7.2	7.5	7.5	7.6	18.9
Self-employment income	16.8	16.2	16.0	16.1	26.5
Supplemental security income (Federal)	7.6	8.4	8.1	8.4	19.9
Social Security	10.8	11.6	11.7	12.3	21.9
AFDC	6.1	6.9	6.5	5.5	16.0
Unemployment compensation	10.1	13.6	10.4	12.7	21.8
Company or union pension	13.9	14.0	12.8	14.7	24.0
Food Stamp allotment	5.2	6.3	6.7	6.6	13.7
Veteran's compensation or pension	11.3	11.2	11.9	13.5	18.3

¹Noninterviews or complete failure to obtain cooperation from any household member have not been considered in this examination of nonresponse rates.

Source: U.S. Bureau of the Census (1985), Report Numbers 3,4,5, and 6, Appendix D.

Comparable rates for SIPP are more difficult to determine. They increase somewhat each quarter because of additional sample loss at each successive wave. Table 5.1 discussed in Section 5.2 shows that 4.9 percent of eligible households were not interviewed in wave 1 of the 1984 panel. McArthur (1987) shows that of sample persons 15 and over who were interviewed in wave 1, the percents not interviewed in the next 3 waves were:

Wave 2	5.6 percent
Wave 3	10.3 percent
Wave 4	14.1 percent

Overall, the comparison favors SIPP, but the advantage of SIPP over the CPS March income supplement in terms of total nonresponse would be considerably smaller in the second calendar year covered by a panel.

5.5 Methodological research and experimentation

This section describes Census Bureau research and experimentation aimed at reducing nonresponse in SIPP and using supplemental data collection procedures to enhance survey content and to compensate for missing data in regular interviews.

EK-3. http://wwwlib.umi.com/dissertations/preview_page/NQ47726/13
İnternet Kaynađı

Chapter 1

Introduction

1.1 Model Assisted Survey Sampling

The last twenty five years has seen a large body of research done with respect to inferential methodologies in finite populations. The thrust of this emphasis is the non applicability of standard methods developed for infinite populations. A finite population may not be satisfactorily characterized by a few parameters, in general, in contrast to infinite populations. However, in practice, since a few parameters such as the mean or variance were of interest, probability mechanisms for generating a sample did provide valid methodology of point estimation (as in Neyman 1934). The basic concern at this juncture was the error of the point estimator, and therefore, concepts such as the bias, standard error, variance, coefficient of variation and mean square error figured prominently in the paradigm. Various sampling mechanisms were developed in order to reduce the sampling variation in the estimator.

Godambe (1955) wrote a monumental paper "*A unified theory of sampling from finite populations*", in which he showed that no '*uniformly best linear unbiased estimator*' exists to estimate a finite population mean or total if one only considers a probability design. Starting with this paper a gradual shift in research activity was initiated. Many concepts that were external to survey sampling were borrowed from the general theory

SEARCH

BROWSE

DIGITAL DISSERTATIONS

ProQuest®

RESULTS LIST | MARKED LIST | SHOPPING CART | DOWNLOAD | CONTACT

24 Page Preview

PRICING ?

Page 13

< Previous Page Thumbnail Index Next Page >

of mathematical statistics and introduced to the sample survey theory. This body of knowledge has come to be known as Superpopulation model theory, Prediction theory and Model assisted designs. The Superpopulation model theory comes in many flavors. Some theoretical survey statisticians (Särndal, 1972, 1976) assume a parametric model for the underlying population and at the same time use the probability mechanism used to gather the sample to estimate the population characteristics under consideration. On the other hand, we have methods advanced by Royall (1968, 1970a, 1970b) that only assume a parametric model for the finite population under consideration and makes no use of the sampling mechanism that generated the sample. Finally, we have Bayesian methods (Ericson, 1969, Rubin, 1983) which also assume a parametric model but make their inferences by using a prior distribution on the parameters. The common theme in all these inferential procedures is the assumption of a parametric model for estimating a population characteristic.

Hansen, Madow and Tepping (1981) presented a paper entitled "Foundations of Inference in Survey Sampling", which has become a corner stone in the survey sampling literature. The thesis of this paper was that only inferences based on probability sampling protected the survey statistician *'against failures of assumed models and provide robustness for all estimators'*. The authors also encouraged research in what they termed model dependent designs. They stated *'The proper use of models has much to contribute to survey design. We urge continuing strong efforts, taking the fullest feasible advantage of models, but ordinarily within the framework of probability sampling, i. e., using designs and estimators that are not model-dependent'*.

1.2 Nonresponse

One of the major problems in all surveys is that of nonresponse. Studies by sampling statisticians, market researchers, and sociologists have attempted to identify methods (sampling and nonsampling) that elicit a high response rate. A unified theory for the

24 Page Preview

PRICING ?

Page 14

<- Previous Page Thumbnail Index Next Page ->

nonresponse problem has not been developed to the same extent as the theory for sampling survey designs.

Nonresponse rates vary widely from survey to survey and from survey organization to survey organization. These rates change over time and even change for repetitions of the same survey. A sample response rate is defined as

$$R = \frac{r}{n}$$

where r is the number of individuals that respond to a survey and n is the number of individuals contacted through the survey. The sample nonresponse rate is defined as

$$R^c = 1 - R.$$

This measure is used to judge the success of a survey; the smaller R^c the better the survey because then the response rate is high. The Panel on Incomplete Data in Sample Surveys (P.I.D.S.S., 1983) came to the following conclusion "*nonresponse and refusal rates have appeared to increase in some surveys, with refusal rates increasing relative to the total nonresponse rates*". As an example, they compared the nonresponse rates for two U. S. federally funded surveys for the years 1968 to 1980. In Table 1 one easily observes that refusal rates have been increasing with time.

Table 1.1
Refusal Rates as a Function of Time

Year	Current Population Survey	Household Health Survey
1968	39%	26%
1969	39%	28%
1970	40%	26%
1971	43%	31%
1972	45%	36%
1973	44%	42%
1974	49%	52%
1975	54%	52%
1976	56%	55%
1977	61%	58%
1978	57%	55%
1979	59%	56%
1980	59%	56%

Source: Published data from the Bureau of the Census for the Current Population Survey and from the National Center for Health for the Household Health Interview Survey.

Many terms and definitions are currently being used to describe various aspects of the non-response problem. The P.I.D.S.S. has addressed this problem and dressed a glossary of non-sampling error terms. We present standardized terminology as developed by P.I.D.S.S.:

Definition 1.2.1 *Non-response refers to missing or incomplete information due to any of the following:*

- a) *the respondent is not available when the data is being collected,*
- b) *the respondent refuses to cooperate with the survey mechanism,*
- c) *the respondent gives a partial answer to questions on the survey.*

24 Page Preview

PRICING ?

Page 16

< Previous Page Thumbnail Index Next Page >

Definition 1.2.2 *Unit nonresponse occurs if a unit is selected for the sample and is eligible for the survey but no response is obtained for the unit or the information obtained is unusable.*

Definition 1.2.3 *Item nonresponse occurs if questions that should be answered are not answered or if answered are classified as unusable. Item nonresponse may be due to any of the following;*

- a) *the respondent does not have the information needed for one or more questions,*
- b) *the respondent refuses to answer a specific question,*
- c) *the interviewer or respondent skip the question.*

1.3 Approaches to the Nonresponse Problem

Statisticians fear that nonrandom response will introduce a severe bias effect in estimates derived from incomplete survey data. One school of thought contends that approaches to the nonrandom response problem should be in the framework of the classical randomization theory traditionally used to analyze sample survey data. The second school of thought argues that the problem would be better handled by finding new methods based on probability models.

Many authors have considered the merits and demerits of each approach and have offered their own suggestions. Bailar, Bailey and Corby (1977) deplore the lack of 'a sound statistical basis for the adjustment procedures' currently being used for analyses of non-response data and welcome 'new simplified methods based on statistical models.' Rubin (1977) uses a Bayesian argument to predict results for nonrespondents given the respondents' data. Methods based on randomization are due to Ernst (1978) and Bailar and Bailar (1978). These authors have used bias adjustments techniques, namely imputation. The two approaches have been combined by Cassel, Sarndal and Wretman (1979). These authors use models to represent the point scatter and then base their analyses of these models on traditional foundations. Griliches, Hall and Hausman (1977), Hausman and

Spence (1977), Little (1983) and Nordheim (1979) have all used probability models for the respondents. These authors regard nonrandom nonresponse as a stochastic censoring process and have used the probit model to represent the censoring mechanism.

1.4 Nonparametric Statistics

While survey statisticians were debating their philosophical foundations, mathematical statisticians were creating a body of knowledge called '*Density estimation, Nonparametric Regression, and General Additive Models*'. All are nonparametric inferential procedures. The basic philosophy as stated by Eubank (1988) for all these inferential procedures is to let '*the data speak for itself*' or as stated by Hastie and Tibshirani (1995) '*let the data show us the appropriate functional form*' without making parametric model assumptions.

Traditional simple regression analysis has its foundations on a known parametric model for the relationship between a response variable y and a predictor variable x . The simplest form of the relationship is the classical linear model

$$y = \alpha + \beta x + \varepsilon,$$

where ε is an error term. The inferential procedure then involves finding estimates of the unknown parameters α and β . The modern regression analysis makes no assumptions on the parametric form of the model. In particular one assumes that

$$y = \mu(x) + \varepsilon,$$

where $\mu(\cdot)$ is an unknown smooth function. The inferential procedure then involves estimating the functional form of the relationship between the response and explanatory variable. The basic idea behind all of these methods is to fit a model to the data points locally. The model will only depend on the observations at a given point and on some specified neighboring points. The fitted model produces estimates of the response variable

ÖZGEÇMİŞ

14.12.1974 tarihinde Elazığ'da doğdu. Orta ve lise öğrenimini T.E.D. Zonguldak Koleji'nde tamamladı. 1998 yılında Hacettepe Üniversitesi Fen Fakültesi İstatistik bölümünü başarıyla tamamladı. 1999-2000 yılları arasında T.C. Devlet İstatistik Enstitüsünde Bilgisayar Programcısı olarak görev yaptı. Gazi Üniversitesi Fen Bilimleri Enstitüsü İstatistik Anabilimdalı Yüksek Lisans programını 2000 yılında kazandı.

