

166138

T.C.

ONDOKUZ MAYIS ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

KLİNİK DENEMELERDE
ÖRNEKLEM BÜYÜKLÜĞÜ BELİRLEMESİ ÜZERİNE
UYGULAMALAR

FERDA ESİN GÜLEL

YÜKSEK LİSANS TEZİ

İSTATİSTİK ANABİLİM DALI

DANIŞMAN

Yrd. Doç. Dr. Mehmet Ali CENGİZ

SAMSUN-2005



KLİNİK DENEMELERDE ÖRNEKLEM
BÜYÜKLÜĞÜ BELİRLEMESİ ÜZERİNE
UYGULAMALAR

FERDA ESİN GÜLEL

YÜKSEK LİSANS TEZİ

İSTATİSTİK ANABİLİM DALI

T.C.
ONDOKUZ MAYIS ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

Bu çalışma, jürimiz tarafından 17/2/2005 tarihinde yapılan sınav ile İstatistik Anabilim Dalı'nda YÜKSEK LİSANS tezi olarak kabul edilmiştir.

Başkan: Prof. Dr. Faruk ALPASLAN

Üye: Yrd. Doç. Dr. Talay AKYILDIZ

Üye: Yrd. Doç. Dr. Mehmet Ali CENGİZ

ONAY:

Yukarıdaki imzaların adı geçen öğretim üyelerine ait olduğunu onaylarım.

09.03/2005



Fen Bilimleri Enstitü Müdürü

Prof. Dr. A. İsmail Öner
Fen Bilimleri Enstitü Müdürü

İÇİNDEKİLER

	Sayfa
ÖZET	i
ABSTRACT.....	ii
TEŞEKKÜR.....	iii
ŞEKİLLERİN LİSTESİ.....	iv
ÇİZELGELERİN LİSTESİ.....	vi
SİMGELER VE KISALTMALAR	vii
1. GİRİŞ	5
2. ÖRNEKLEM GENİŞLİĞİNİN BELİRLENMESİ	6
2.1. Tek Örneklem Durumunda Örneklem Genişliği Belirleme	6
2.1.1. Normal Dağılımlı Bir Anakütleden Örnek Seçimi	6
σ^2 Bilindiğinde μ Tahmini	7
σ^2 Bilinmediğinde μ Tahmini	8
σ^2 'nin Tahmini	8
σ^2 Bilindiğinde μ Hakkında Hipotez Testleri	12
σ^2 Bilinmediğinde μ Hakkında Hipotez Testleri	14
σ^2 Hakkında Hipotez Testleri	16
2.1.2. Bernoulli Dağılımlı Bir Anakütleden Örnek Seçimi	18
θ 'nın Tahmini	18
θ Hakkında Hipotez Testi	20
2.1.3. Tahmin için Bayes Yaklaşımı	22
2.2. İki Örneklem Durumunda Örneklem Genişliği Belirleme	23

2.2.1.	Normal Dağılımlı İki Anakütleden Örnek Seçimi	24
	σ_1^2 ve σ_2^2 Bilindiğinde $\mu_1 - \mu_2$ Tahmini	24
	$\sigma_1^2 = \sigma_2^2$ Olduğunda ve Ortak Varyans Bilinmediğinde $\mu_1 - \mu_2$ Tahmini	26
	$\sigma_1^2 \neq \sigma_2^2$ Olduğunda ve Varyanslar Bilinmediğinde $\mu_1 - \mu_2$ Tahmini	26
	σ_1^2 / σ_2^2 Varyans Oranının Tahmini	27
	σ_1^2 ve σ_2^2 Bilindiğinde μ_1 ve μ_2 ile İlgili Hipotez Testi	30
	$\sigma_1^2 = \sigma_2^2$ Olduğunda ve Ortak Varyans Bilinmediğinde μ_1 ve μ_2 ile İlgili Hipotez Testi	31
	$\sigma_1^2 \neq \sigma_2^2$ Olduğunda ve Varyanslar Bilinmediğinde μ_1 ve μ_2 ile İlgili Hipotez Testi	33
	σ_1^2 ve σ_2^2 ile İlgili Hipotez Testi	35
2.2.2.	Bernoulli Dağılımlı İki Anakütleden Örnek Seçimi	36
2.3.	K Örneklem Durumunda Örneklem Genişliği Belirleme	37
2.3.1.	Ortalamalar Farkının Tahmini	38
	Dağılıma Ait Varsayımlar Olmadan Tahmin	38
	Normallik Varsayımı Altında Tahmin	40
2.3.2.	Normal Dağılımlı Bir Anakütleden Örnek Seçimi	42
2.3.3.	Bernoulli Dağılımlı Bir Anakütleden Örnek Seçimi	44
2.4.	Sonlu Anakütleden Örnek Seçimi	47
2.4.1.	Basit Rastgele Örneklem	47
	Kabul Örnekleme	49
2.4.2.	Tabakalı Rastgele Örnekleme	50
2.4.3.	Tabakalamaya Göre Tam Örnekleme	52

2.4.4. İki Aşamalı Küme Örneklemesi	53
3. KLİNİK ÇALIŞMALARDA ÖRNEKLEM SEÇİMİ	56
3.1. Nitel Bağımlı Değişkenlerle Deney	56
3.1.1. McNemar Testi	56
3.1.2. Kategorik Bağımlılarla (c) Tedavi Etkilerinin (r) Eşitliği Testi	61
3.1.3. Uzun Dönem Denemeler	62
3.1.4. Olgu–Kontrol Çalışmaları	63
3.1.5. İki Aşamalı Tasarım	66
3.2. Nicel Bağımlı Değişkenlerle Deney	66
3.2.1. Biyoeşdeğerliğe Göre Sınama	66
3.2.2. İki Süreç, İki Tedavi, Çapraz Çalışması	70
3.2.3. Yaşam Dağılımlarına Benzerliğine Göre Klinik Denemeler	71
3.3. Lojistik Regresyon için Örneklem Genişliği ve Güç	74
3.3.1. İkili Oran Karşılaştırması için Örneklem Genişliği	74
3.3.2. Nicel (Sayısal) Tahmin Edicilerle Örneklem Genişliği	75
3.3.3. Çoklu Lojistik Regresyonda Örneklem Genişliği	75
4. UYGULAMA	77
4.1. Anakütle Ortalamasının Tahmininde Örneklem Genişliği Belirleme	77
4.2. Ortalamalar Arasındaki Farklılığın Karşılaştırılmasında Örneklem Genişliği Belirleme	79
4.3. Anakütle Oranının Tahmininde Örneklem Genişliği Belirleme	81
4.4. Oranlar Arasındaki Farklılığın Karşılaştırılmasında Örneklem Genişliği Belirleme	84
4.5. Bir Anakütlede Hastalığı Teşhis Etmek için Gereken Örneklem Genişliğinin Belirlenmesi	86

4.6. Eşleştirilmemiş Olgu–Kontrol Çalışmalarında Örneklem Genişliği Belirleme	87
4.7. Eşleştirilmiş Olgu–Kontrol Çalışmalarında Örneklem Genişliği Belirleme	88
4.8. Kohort (Grup İzleme) Çalışmalarında Örneklem Genişliği Belirleme	89
4.9. Lojistik Regresyonda Örneklem Genişliği Belirleme	90
SONUÇ VE ÖNERİLER.....	92
KAYNAKLAR.....	93
ÖZGEÇMİŞ.....	98



KLİNİK DENEMELERDE ÖRNEKLEM GENİŞLİĞİ BELİRLEMESİ ÜZERİNE UYGULAMALAR ÖZET

İstatistiksel çalışmalar(anketler, deneyler ve gözlemsel çalışmalar), dikkatli planlandığında daima daha iyidir. İyi planlamanın bir çok yönü vardır. Problem dikkatlice tanımlanmalıdır. Deneysel veya gözlemsel birimler uygun anakütleden seçilmeli ve çalışma doğru bir şekilde rastgeleştirilmelidir. İşlemler dikkatli takip edilmeli ve sonuçları elde etmek için güvenilir araçlar kullanılmalıdır. Sonuçta çalışmanın amacına uygun olarak yeterli örnek büyüklüğü seçilmelidir.

Klinik çalışmalarda çalışmaya katılacak birey sayısı yani çalışmanın örnek büyüklüğü, önemli bir problemdir. Uygulamalı istatistikçiler ve istatistik kullanıcıları, sık olarak “Örneklem büyüklüğünü nasıl seçmeliyim?” sorusunu sorarlar. Basit rastsal bir örneğin uygun olacağı kabul edilse bile bu sorunun cevaplanması kolay değildir.

Böyle bir örnek büyüklüğünün etkisinin bilimsel olarak anlamlılığı, aynı zamanda istatistiksel olarak anlamlı olacağının ifadesidir. Bilimsel olarak yeterince büyük olması her zaman istatistiksel olarak büyüklük anlamına gelmez. Örnek büyüklüğü, ekonomik nedenler açısından da önemlidir. Örneklem büyüklüğü yeterli olandan azsa, doğru sonuçlar elde edilemeyeceğinden kaynakların gereksiz kullanımına neden olurken gereğinden büyük örnek seçimi de kaynakların aşırı olarak israf edilmesine neden olur. İnsan ve hayvan denekleri içeren denemelerde örnek seçiminin etik nedenleri de vardır. Düşük dozda ilaç kullanımı tedaviyi tam yapamayacak ya da değişik yan etkilere yol açacaktır. Aşırı doz kullanımı ise faydadan çok potansiyel bir zarar verecektir.

Bu çalışma, uygun örnek büyüklüğü için temel faktörleri ve hesaplamaları özetleyerek bazı medikal problemler için Win Episcopy 2.0, MINITAB ve S-Plus 2000 gibi paket programlardaki çözümleri verecektir.

Anahtar Kelimeler: Klinik Denemeler, Örnek Büyüklüğü, Tahmin

APPLICATIONS ON SAMPLE SIZE ESTIMATION IN CLINICAL TRIALS

ABSTRACT

Statistical studies (surveys, experiments, observational studies, etc.) are always better when they are carefully planned. Good planning has many aspects. The problem should be carefully defined. Experimental or observational units must be selected from the appropriate population. The study must be randomized correctly. The procedures must be followed carefully. Reliable instruments should be used to obtain measurements. Finally, the study must be of adequate size, relative to the goals of the study.

The number of individuals to include in a research study, the sample size of the study, is an important consideration in the design of many clinical studies. Practicing statisticians and the users of statistics are often asked 'How big a sample should I take?'. Even if it is accepted that a simple random sample is appropriate, the question is often not an easy one to answer.

It must be "big enough" that an effect of such magnitude as to be of scientific significance will also be statistically significant. It is just as important, however, that the study not be "too big," where an effect of little scientific importance is nevertheless statistically detectable. Sample size is important for economic reasons: An under-sized study can be a waste of resources for not having the capability to produce useful results, while an over-sized one uses more resources than are necessary. In an experiment involving human or animal subjects, sample size is a pivotal issue for ethical reasons. An under-sized experiment exposes the subjects to potentially harmful treatments without advancing knowledge. In an over-sized experiment, an unnecessary number of subjects are exposed to a potentially harmful treatment, or are denied a potentially beneficial one.

This study reviews the basic factors that determine an appropriate sample size and provides methods for its calculation in some simple cases gives the results of some medical problems using some pocket programs such as Win Episcope 2.0, Minitab and S-Plus 2000.

Key Words: Clinical Trials, Sample Size, Estimation

TEŞEKKÜR

Bu çalışmada bana destek veren danışmanım Yrd. Doç. Dr. Mehmet Ali CENGİZ'e yardımlarından dolayı teşekkür ederim.

Babama, anneme ve abime; her konuda bana destek oldukları için,

Nurdan'a, Özgür'e, Zeliha'ya ve Yeliz'e; bana "yapabileceğim" konusunda moral verdikleri için,

Aliye ve Raşit SARIKAYA'ya; bana yol gösterdikleri için,

Araş. Gör. Demet HAN'a; sabrı için,

Ve SİNOP'a; bu çalışma sırasında beni hırslandırdığı için

Teşekkür ederim.



ŞEKİLLERİN LİSTESİ

	Sayfa
4.1. Anakütle Ortalamasının Tahmininde Örneklem Genişliği Belirleme Win Episcopo 2.0 Çözümü	77
4.2. Anakütle Ortalamasının Tahmininde Örneklem Genişliği Belirleme Minitab Çözümü	78
4.3. Anakütle Ortalamasının Tahmininde Örneklem Genişliği Belirleme S-Plus 2000 Çözümü.....	78
4.4. Ortalamalar Arasındaki Farklılığın Karşılaştırılmasında Örneklem Genişliği Belirleme Win Episcopo 2.0 Çözümü.....	79
4.5. Ortalamalar Arasındaki Farklılığın Karşılaştırılmasında Örneklem Genişliği Belirleme Minitab Çözümü.....	80
4.6. Ortalamalar Arasındaki Farklılığın Karşılaştırılmasında Örneklem Genişliği Belirleme S-Plus 2000 Çözümü	80
4.7. Anakütle Oranının Tahmininde Örneklem Genişliği Belirleme Win Episcopo 2.0 Çözümü	81
4.8. Anakütle Oranının Tahmininde Örneklem Genişliği Belirleme Minitab Çözümü	82
4.9. Anakütle Oranının Tahmininde Örneklem Genişliği Belirleme S-Plus 2000 Çözümü.....	83
4.10. Oranlar Arasındaki Farklılığın Karşılaştırılmasında Örneklem Genişliği Belirleme Win Episcopo 2.0 Çözümü.....	84
4.11. Oranlar Arasındaki Farklılığın Karşılaştırılmasında Örneklem Genişliği Belirleme Minitab Çözümü.....	85
4.12. Oranlar Arasındaki Farklılığın Karşılaştırılmasında Örneklem Genişliği Belirleme S-Plus 2000 Çözümü	85
4.13. Bir Anaküttelede Hastalığı Teşhis Etmek İçin Gereken Örneklem Genişliğinin Belirlenmesi Win Episcopo 2.0 Çözümü	86
4.14. Eşleştirilmemiş Olgu-Kontrol Çalışmalarında Örneklem Genişliği Belirleme Win Episcopo 2.0 Çözümü.....	87
4.15. Eşleştirilmiş Olgu Kontrol Çalışmalarında Örneklem Genişliği Belirleme Win Episcopo 2.0 Çözümü.....	88

4.16. Kohort (Grup İzleme) Çalışmalarında**Örneklem Genişliği Belirleme Win Episcpe 2.0 Çözümü..... 89**

ÇİZELGELERİN LİSTESİ

	Sayfa
3.1. Tedavi A ve B'ye Göre Birimlerin Sayısı	57
3.2. Bağımlı Kategoriler ve Tedavilere Göre Birimlerin Sayısı	61
3.3. Etki Karışımı Değişkeninin Derecesine Dayanan j. Tabakadaki Veriler	63
3.4. Birbirini İzleyen Tedavilerde Deneysel Düzen	70
4.1. Lojistik regresyonda çeşitli α ve β değerlerine göre Örneklem genişliği hesaplamaları	91



SİMGELER ve KISALTMALAR

c^*	:	Belirli Toplam Beklenen Maliyet
cs	:	Correct Selection (Doğru Seçim)
$D(.,.,.,.,.)$	:	İddiaların Diyagonal (Köşegen) Matrisi
e_{ijk}	:	Rastgele Hata
h	:	Bechoffer v.d.(1954) Tarafından Hazırlanan Tablo Değeri
$J_{k-1,k-1}$	:	Tamamıyla Aynı $(k-1) \times (k-1)$ Matrisi
n_{ij}	:	i. Tedaviden Alınan j. Bağımlı Kategorili Eleman
$[n^*]$	:	En Büyük Tam Sayı Fonksiyonu
N_h	:	Anakütlenin Toplam Genişliği
s_h^2	:	h. Tabaka Örneklem Varyansı ($h = 1, 2, \dots, L$)
S_j	:	j. Tabakadaki Olguların Oranı
S^2	:	N Genişlikli Anakütlenin Varyansı
S_p^2	:	Birleştirilmiş Varyans
$S_{w_i}^2$	:	i. Kümede Bağımlı İkinci Aşama Birimlerin Varyansı
t_j	:	j. Tabakanın Nisbi Genişliği
$t_{p,v}$	:	v Serbestlik Dereceli Student t Tablo Değeri
		Sayısı ($i = 1, 2, \dots, r; j = 1, 2, \dots, c$)
Y_i	:	i. Kümedeki Toplam Bağımlı
Y_{ijk}	:	i. Etkide, j. Denekte, K. Periyottaki
$\bar{Y}$	:	Küme Ortalama Bağımlısı
$\bar{Y}_i$	:	i. Kümede Bağımlı İkinci Aşama Birimlerinin Ortalaması
$\hat{\bar{X}}$	:	$\bar{X}$ 'nın Yansız Tahmin Edicisi
$Z_{\alpha/2}$	:	Standart Normal Dağılım Tablo Değeri
$\Phi(.)$	:	Standart Normal Değişkenin Dağılım Fonksiyonu
μ	:	Genel Ortalama
ρ_i	:	i. Periyod Etkisi

π_{1j} ve π_{2j}	:	j. Tabakada ($j = 1, 2, \dots, J$) sırasıyla Olgu ve Kontrol Gruplarındaki Riske Göre Etkilenenlerin Olasılıkları Bağımlılık ($i = 1, 2; j = 1, 2, \dots, n; k = 1, 2$)
τ_i	:	i. Tedavinin Doğrudan Etkisi
δ_i	:	i. Tedavinin Artık Etkisi
γ_{ij}	:	i. Etkide j. Denek Etkisi



1. GİRİŞ

İstatistiksel çalışmalar(anketler, deneyler, gözlemsel çalışmalar, vs.) dikkatlice planlandıklarında iyi sonuçlar elde edilir. Bu çalışmalarda problem dikkatli bir şekilde planlanmalı ve işlemsel duruma getirilmelidir. Deneysel veya gözlemsel birimler, uygun anakütleden seçilmeli ve çalışılacak örneklemin genişliği doğru belirlenmelidir. Daha iyi sonuç elde etmek için genellikle büyük örneklemlemlerle çalışmak tercih edilir. Ancak bu, zaman, para ve emek kaybına neden olur. Küçük örneklemlemlerle çalışıldığında ise örneklemden yapılacak tahminler gerçeği ortaya koymayabilir. Bu nedenle yeterli genişlikte örnek seçmek ve seçilen örneklemin genişliğinin uygun olduğunu söyleyebilmek gerekir.

Örneklem genişliği; maliyet, anakütle parametresi tahmininin standart hatası, anakütle parametresi ile tahmini arasındaki fark (mutlak hata), OC (işletim karakteristik) eğrileri veya hazırlanmış tablolar kullanılarak belirlenebilir. Ama en yaygın belirleme ; güç hesaplamasından ya da güven aralığı genişliğinden yapılır. Güç yaklaşımı aşağıda verilen öğeleri içerir:

- i) θ parametresi ile veriler için olasılık modelinin temelini oluşturulmasıyla hipotez testini belirlemek
- ii) Testin α önemlilik düzeyini belirlemek
- iii) Bilimsel olarak ilgilenilen alternatifi yansıtan $\tilde{\theta}$ etki genişliğini belirlemek
- iv) Testin güç fonksiyonunu hesaplamak için gereken diğer parametrelerin tahminini veya kayıtlı verileri tespit etmek
- v) $\theta = \tilde{\theta}$ olduğunda testin gücünün $\tilde{\pi}$ hedef değerini belirlemek

Testin gücü; $\pi(\theta, n, \alpha, \dots)$ fonksiyonudur. n ; örneklem genişliği ve "... " kısmı incelenen anakütlerde ilgilenilen parametrelerdir. Gereken örneklem genişliği;

$$\pi(\tilde{\theta}, n, \alpha, \dots) \geq \tilde{\pi}$$

eşitsizliğini sağlayan en küçük tamsayıdır. Bazen güç fonksiyonunda, diğer parametreleri ve varyans tahminlerini kullanabilecek kayıtlı veriler olur. Eğer böyle bir kayıt yoksa pilot çalışma yapmak gerekir.

Güç hesaplaması ve güven aralığı yöntemlerinden belirlenen örneklem genişliklerinin şekilleri benzer olurken hedefleri farklıdır ve sıklıkla çok farklı örneklem

genişliği bağıntılarına neden olurlar. Diğer bir örneklem genişliğini belirlemede kullanılan yöntem; örneklem genişliği ile tahmin edicinin standart hatası arasındaki ilişkiden meydana gelir. Bu yöntemde kullanılan formülün seçilen girdilere yüksek duyarlılıkta olabilmesi için hedef kriterinin ve parametre tahminlerinin dikkatli seçimi kesin örneklem genişliğini belirlemede temel adımdır. Yani asıl konu, belirli bir hata payı ile anakütle parametrelerinin tahmin edilmesidir. Örneklem genişliği ile hata payı arasında ters bir ilişki vardır. Örneklem genişliği arttığında hata payı küçülür ama maliyet artar. Bu ifadenin tersi de geçerlidir. Uygun örneklem genişliği; genişlik, hata payı ve maliyet optimum bir dengede tutulduğunda belirlenir. Zaten araştırmacının istediği örneklem genişliği de en küçük maliyetle önceden belirlenen hata sınırını aşmamak koşuluyla tahminler yapabilmeyi sağlayan genişliktir.

Örnek büyüklüğünün belirlenmesi, bir klinik çalışmanın tasarım aşamasında karşılaşılan önemli problemlerden biridir. Bu problem, çalışma hipotezini belirleyecek istatistiksel işlemler ve tasarım ile doğrudan bağlantılıdır. Çalışma hipotezi; karşılaştırılacak tedavi türleri, her tedavi grubundaki hastaların sayısı ve nitelikleri gibi parametrelerin belirlenmesinde yardımcı olur. Çalışma için ne kadar hastaya (deneğe) yani örnek büyüklüğüne ihtiyaç olduğu bu parametrelerden ilk akla gelendir.

Örneklem genişliği belirleme bir çalışma için en önemli aşamalardan biri olmasına rağmen bu konu hakkında şaşırtıcı derecede az çalışma yapılmıştır. Önemli referanslar olarak Mace(1964), Kraemer ve Thiemann(1987), Cohen(1988), Odeh ve Fox(1991) verilebilir. Özellikle biyoistatistik dergilerinde, belirli testler için örneklem genişliği belirleme konusuyla ilgili çalışmalar bulunmaktadır. İlgilenilen şeylerden biri de örneklem genişliğinin yeterli veya yetersiz olmasının dışında örneklem genişliğinin boyutudur. Bu konuyla ilgili Freiman v.d.(1986) ve Thornley ve Adams(1998) referans olarak gösterilebilir.

Klinik denemelerin yaşam dağılımlarına benzerlik gösterdiği durumda örneklem genişliğini belirlemek için Rubinstein v.d.(1981) bir yöntem geliştirmiş, George ve Desu(1974) bu yönetime alternatif olarak gösterilmiştir.

Schoenfeld ve Richter(1981), iki tedavinin karşılaştırıldığı ve yaşamın son nokta olduğu klinik denemeler için hesaplanan örneklem genişliklerine göre (gereken hasta sayısı) nomogram hazırladı. Bu hazırlanan nomogramlarda, yaşam zamanının üstel dağılıma uyduğu ve hastaların çalışmaya tekdüze girdiği varsayıldı.

Schoenfeld(1981), verinin, iki yaşam dağılımını karşılaştırmak için kullanılan Logrank testiyle analiz edilmesi durumunda örneklem genişliği belirleme problemini inceledi. Aynı problemi Freedman(1982) da inceledi ve bu çalışma için gereken örneklem genişliği sayısını tablolastırdı.

Bazı klinik denemelerde üç veya daha fazla tedavi grubunu karşılaştırmak amaçlanır. Makuch ve Simon(1982), k tedavi grubu arasında başarısızlık zamanını (yani ölüm zamanını) karşılaştırdığında, deneme süresi ve örneklem genişliği koşulunu belirleme problemini inceledi.

Simülasyon çalışmasında Rubinstein v.d.(1981); “Klinik denemelerin yaşam dağılımlarına benzerlik gösterdiği durumda örneklem genişliğini belirlemek için kullanılan yöntem, parametrik olmayan Mantel–Haenszel(logrank) testi ile uygulandığında doğru gücü veren deneme süresini elde eder.” görüşünü ileri sürdü. Takibe göre kayıp ve takip dönemi olmadığında deneme süresi; George ve Desu(1974)’nın bulduğu sonuçlardan çok küçük farkla Rubinstein v.d.(1981) tarafından elde edildi.

Morgan(1985), klinik denemelerde iyileşme periyodunun seçimine ve belirli farkları ortaya çıkarmak için yeterli güç üretiminde takip periyodunun tek olmadığına dikkat çekti. İyileşme ve takip periyodları için uygun kombinasyonların belirtimine, takip ve iyileşme dönemlerinin sadece maliyet etkisi seçimine yöntemler sundu.

Takibin maliyet etkisi sürekliliğinin, son noktanın belirlenmesine göre takip edilen hastaların maliyetinin oranına bağlı olduğundan maliyeti minimize edilen klinik denemelerin en uygun süresi Gross v.d.(1987) tarafından da incelenmiştir.

Palta ve Amini(1985) tarafından, koşullar dikkate alınarak, yaşam çalışmaları için örneklem genişliği belirtimi incelendi, çok merkezli klinik deneme durumlarına uygulanan formüller ve tablolar oluşturuldu. Buradan; ortak değişken, tabakalarda gruplandığı zaman örneklem genişliği tahmini de elde edilebilir.

Genişlik belirlemede kolaylık sağlayan yazılımlar da geliştirilmiştir. Bunlardan bazıları; UnifyPow(O’Brien, 1998), Power and Precision(Borenstein, 1997), nQuery Advisor(Elashoff, 2000), PASS(Hintze, 2000)’dır. Web kaynakları, güç analiz yazılımının (Thomas, 1998) kapsamlı listesini ve Lenth(2000) gibi çevrim içi hesaplamaları içerir. Wheeler(1974), doğrusal modellerde kullanım için bazı yararlı

yaklaşımları geliştirmiştir. Castelloe(2000), hesap yöntemlerinin güncellenmiş genel açıklamalarını verir.

Örneklem genişliğini belirlerken Bayes Yöntemi de kullanılabilir. Bu yöntem de bilinmeyen parametreler için Bayesci önsel dağılımı kullanılarak, beklenen maliyet ve faydaların ayarlanmasıyla örneklem genişliği belirlenebilir. Bu yöntemle ilgili Grundy v.d.(1956)'nin makalesi incelenebilir. Bayesyen yöntemin metodolojisi detaylı olarak Raiffa ve Schlaifer(1961) tarafından verilmiştir ve son zamanlarda Lindley(1997) tarafından tanımlanmıştır.

Örnek büyüklüğünün tahmini, denemenin istatistiksel anlamlılığı açısından oldukça önemli olmasına ve klinik çalışmalarda örnek büyüklüğünün tahmini için kullanılan birçok paket program bulunmasına rağmen ülkemizde pek çok alanda olduğu gibi klinik çalışmalarda da çoğu zaman bu faktör göz ardı edilmektedir.

Örneklem genişliğinin belirlenmesi için gerekli olan yöntemlerin tartışıldığı bu çalışmanın;

Birinci bölümünde; bir çalışmada örneklem genişliği belirlemenin önemi ve bu konu ile ilgili yapılan çalışmalar sunulmuştur.

İkinci bölümünde; tek örneklem, iki örneklem, k örneklem durumlarında örneklem genişliği belirleme yöntemleri verilmiştir. Tek örneklem ve iki örneklem durumlarında örneklem genişliği belirleme Normal dağılımlı ve Bernoulli dağılımlı anakütlelerden örnek seçimi olarak iki başlık altında incelenmiş ve bu başlıklarda Normal dağılımlı anakütleden örnek seçiminde anakütle varyansının bilinip bilinmemesine bağlı olarak ortalama tahmini, ortalama hakkında hipotez testleri ve varyans tahmini, varyans hakkında hipotez testleri; Bernoulli dağılımlı bir anakütleden örnek seçimi durumunda ise θ tahmini ve θ hakkında hipotez testleri olarak ayrıntılı bilgi verilmiştir. K örneklem durumunda ise ortalamalar farkının tahmini, normal ve Bernoulli dağılımlı bir anakütleden örnek seçimi, sonlu anakütleden örnek seçimi, sıralama-seçim başlıkları verilmiştir. Ortalamalar farkının tahmini başlığında; dağılıma ait varsayımlar olmadan tahmin ve normallik varsayımı altında tahmin alt başlıklarıyla örneklem genişliği belirleme ayrıntılı olarak incelenmiştir. Sonlu anakütleden örnek seçimi başlığı; kabul örneklemesi, tabakalı rastgele örneklem, tabakalamaya göre tam örneklem, iki aşamalı küme örneklemesi alt başlıklarıyla verilmiştir.

Üçüncü bölümünde; klinik çalışmalarda örneklem seçimi ana başlığı altında Nitel bağımlı değişkenlerle deney başlığında; McNemar testi, Kategorik bağımlılarla tedavi etkilerinin eşitliği testi, uzun dönem denemeler, olgu-kontrol çalışmaları, iki aşamalı tasarım, nicel bağımlı değişkenlerle deney başlığında, biyoeşdeğerliğe göre sınaama, çapraz çalışması, yaşam dağılımlarına benzerliğine göre klinik denemeler ve lojistik regresyon için örneklem genişliği ve güç başlığında, ikili oran karşılaştırmasında örneklem genişliği belirleme, nicel tahmin edicilerle örneklem genişliği belirleme ve çoklu lojistik regresyonda örneklem genişliği belirleme konuları alt başlıklar halinde verilmiştir.

Dördüncü bölümünde; ikinci ve üçüncü bölümlerde incelenen konulardan bazıları ile ilgili Win Episcpe 2.0, Minitab ve S-Plus 2000 uygulamaları verilmiştir.



2. ÖRNEKLEM GENİŞLİĞİNİN BELİRLENMESİ

2.1. Tek Örneklem Durumunda Örneklem Genişliği Belirleme

Bu bölümde, belirli olasılık dağılımlarından çekilen rastsal örneklemelerin genişliklerini belirleme yöntemleri incelenmiştir.

2.1.1. Normal Dağılımlı Bir Anakütleden Örnek Seçimi

X sürekli rastlantı değişkeninin; ortalaması μ , varyansı σ^2 olan normal dağılıma sahip olduğunu varsayalım. Yani $X \sim N(\mu, \sigma^2)$ ve dağılımın olasılık yoğunluk fonksiyonu

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, -\infty < x < \infty \text{ 'dır.}$$

Normal dağılımlı $X_1, X_2, \dots, X_n$ rastlantı değişkenlerinden oluşan X rastsal örnekleme ele alınırsa, X'in ortalama ve varyansı

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n} \quad S^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{(n-1)} \text{ 'dır.}$$

Bu çalışmanın ileriki bölümlerinde kullanılacak Normal dağılımın temel özellikleri aşağıda verilmiştir:

1. Ortalaması μ , varyansı σ^2 olan normal bir anakütleden çekilmiş n büyüklüğünde rastsal örneklemin ortalaması $\bar{X}$ ise; $\bar{X}$ 'nın örnekleme dağılımı da ortalaması μ , varyansı σ^2/n olan bir normal dağılımdır.
2. Ortalaması μ , varyansı σ^2 olan normal bir anakütleden çekilmiş n büyüklüğünde rastsal örneklemin ortalaması $\bar{X}$ ve varyansı S^2 ise,
 - (i) $\bar{X}$ ile S^2 bağımsızdır.
 - (ii) $\frac{(n-1)S^2}{\sigma^2}$ rastlantı değişkeni, $(n-1)$ serbestlik derecesiyle χ^2 (ki - kare) dağılır.

Örneklem varyansının olasılık yoğunluk fonksiyonu aşağıdaki gibidir:

$$f(y; \nu) = \frac{1}{\Gamma\left(\frac{\nu}{2}\right) 2^{\frac{\nu}{2}}} y^{\frac{\nu}{2}-1} e^{-\frac{y}{2}}, 0 < y < \infty$$

3. Ortalaması μ , varyansı σ^2 olan normal bir anakütleden çekilmiş n büyüklüğünde rastsal örneklemin ortalaması $\bar{X}$ ve varyansı S^2 ise,

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}}$$

değişkeni, $(n-1)$ serbestlik derecesiyle t dağılır. Olasılık yoğunluk fonksiyonu

$$f(t; \nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi\nu} \Gamma\left(\frac{\nu}{2}\right)} \cdot \frac{1}{\left(1 + \frac{t^2}{\nu}\right)^{(\nu+1)/2}} \text{ 'dır.}$$

μ ve σ^2 tahminleri ve bu parametrelere ilişkin hipotez testleri incelenecektir.

σ^2 bilindiğinde μ tahmini

μ 'nün en iyi yansız tahmin edicisi $\bar{x}$ 'dır. μ tahmininde Mutlak Hata; $|\bar{x} - \mu|$ 'ye eşittir.

Teorem 2.1: d ve α önceden belirlenmiş sabitler olmak üzere

$P\left(|\bar{x} - \mu| \leq d\right) \geq 1 - \alpha$ koşulunu sağlayan örneklem genişliği; $[n^*] + 1$ 'dir ve

$n^* = \left(Z_{\alpha/2} \cdot \frac{\sigma}{d}\right)^2$ 'dir. Burada $[n^*]$, en büyük tamsayıyı veren fonksiyondur.

İspat:

$$P\left(|\bar{x} - \mu| \leq d\right) \geq 1 - \alpha$$

$$P\left(\left|\frac{\bar{x} - \mu}{\sigma/\sqrt{n}}\right| \leq \frac{d}{\sigma/\sqrt{n}}\right) \geq 1 - \alpha$$

$$P\left(Z \leq \frac{d}{\sigma/\sqrt{n}}\right) \geq 1 - \alpha$$

$$Z_{\alpha/2} = \frac{d}{\sigma / \sqrt{n}}$$

$$n^* = \left(Z_{\alpha/2} \cdot \frac{\sigma}{d} \right)^2$$

σ^2 bilinmediğinde μ 'nün tahmini

σ^2 bilinmediğinde μ tahmin edilirken

$$P \left(\left| \bar{x} - \mu \right| \leq d \right) \geq 1 - \alpha$$

koşulunda d ; σ^2 'nin eşiti olarak yazılır. $d = k\sigma$ ise örneklem genişliği, k sabit bir sayı olmak üzere aşağıdaki formülden hesaplanır:

$$n^* = \left(Z_{\alpha/2} \cdot \frac{\sigma}{k\sigma} \right)^2 = \left(Z_{\alpha/2} \cdot \frac{1}{k} \right)^2$$

μ 'yü tahmin etmek için $d = k\sigma$ eşitliğinin yazılamadığı durumda Stein(1945)'in iki aşamalı metodu uygulanır:

1. İlk olarak n_1 genişlikli, S_1^2 varyanslı örneklem alınır.
2. $n = \max \left\{ n_1, \left\lceil \frac{S_1^2 \cdot t_{\alpha/2, (n_1-1)}^2}{d^2} \right\rceil + 1 \right\}$
3. Gerekirse $(n - n_1)$ kadar daha gözlem alınır.
4. n tane gözlemin ortalaması $\bar{x}$, $P \left(\left| \bar{x} - \mu \right| \leq d \right) \geq 1 - \alpha$ koşulunu sağlar.

İki aşamalı yöntemde n_1 ilk örneklem genişliğinin iyi seçilmesi gerekir. Farklı n_1 ve σ^2 değerleri için $E(n)$ alınarak n_1 'in en uygun seçimi yapılır(Seelbinder, 1953; Moshman, 1958)

σ^2 'nin tahmini

μ 'nün bilinip bilinmemesine göre anakütle varyansının tahmin edicileri aşağıdaki gibidir:

$$\mu \text{ biliniyorsa ; } \hat{\sigma}^2 = \sum_{i=1}^n \frac{(X_i - \mu)^2}{n}$$

μ bilinmiyorsa ; S^2

Örneklem varyansı; μ biliniyorsa $\frac{n\hat{\sigma}^2}{\sigma^2}$, n serbestlik derecesiyle χ^2 , μ bilinmiyorsa $\frac{(n-1)\hat{\sigma}^2}{\sigma^2}$, $(n-1)$ serbestlik derecesiyle χ^2 dağılır.

σ^2 tahmininde örneklem genişliği belirlemek için σ^2 hakkında önsel bir bilgiye sahip olmamız gerekir. Bu önsel bilgi olmadan

$$P\left(\left|\hat{\sigma}^2 - \sigma^2\right| \leq d\right)$$

koşulu sağlanamaz. Dolayısıyla örneklem genişliği de belirlenemez. Ancak $d = r\sigma^2$ alınırsa,

$$P\left(\left|\hat{\sigma}^2 - \sigma^2\right| \leq r\sigma^2\right) = P\left(\left|\hat{\sigma}^2 - \sigma^2\right| / \sigma^2 \leq r\right)$$

koşulu ile belirlenebilir. Koşulun değiştirilmesiyle

$$\left|\frac{\hat{\sigma}^2 - \sigma^2}{\sigma^2}\right|$$

nisbi hata kontrol edilir. Böylece n örneklem genişliği,

$$P\left(\left|\frac{\hat{\sigma}^2 - \sigma^2}{\sigma^2}\right| \leq r\right) \geq 1 - \alpha \quad (2.1)$$

koşulunun sağlanmasıyla bulunur.

$$P\left[(1-r)f \leq \frac{f\hat{\sigma}^2}{\sigma} \leq (1+r)f\right] \geq 1 - \alpha \quad (2.2) \quad \mu \text{ biliniyorsa; } f = n$$

μ bilinmiyorsa; $f = n-1$ 'dir.

$G_f(\cdot)$, f serbestlik dereceli χ^2 değişkeninin dağılım fonksiyonu olsun. (2.2) eşitliği,

$$G_f((1+r)f) - G_f((1-r)f) \geq 1 - \alpha \quad (2.3)$$

olarak yazılabilir.

Teorem 2.2: $P\left(\left|\frac{\hat{\sigma}^2 - \sigma^2}{\sigma^2}\right| \leq r\right) \geq 1 - \alpha$ koşulunu sağlamak için gereken örneklem genişliği n ; μ biliniyorsa f , μ bilinmiyorsa $(f+1)$ 'dir. f ; (2.3) koşulunu sağlayan en küçük pozitif tam sayıdır.

İspat:

$$P\left(\left|\frac{\hat{\sigma}^2 - \sigma^2}{\sigma^2}\right| \leq r\right) \geq 1 - \alpha$$

$$P\left(-r \leq \frac{\hat{\sigma}^2 - \sigma^2}{\sigma^2} \leq r\right) \geq 1 - \alpha$$

$$P\left(-r \leq \frac{\hat{\sigma}^2}{\sigma^2} - 1 \leq r\right) \geq 1 - \alpha$$

$$P\left(1 - r \leq \frac{\hat{\sigma}^2}{\sigma^2} \leq 1 + r\right) \geq 1 - \alpha$$

$$P\left((1-r) \cdot f \leq \frac{f \cdot \hat{\sigma}^2}{\sigma^2} \leq (1+r) \cdot f\right) \geq 1 - \alpha \rightarrow \chi_f^2 \text{ dağılımı}$$

$$G_f((1+r)f) - G_f((1-r)f) \geq 1 - \alpha$$

f ; (2.3) eşitliğinden kolayca elde edilemez. f serbestlik derecesini elde etmek için kaba bir yaklaşım şu şekildedir:

Merkezi Limit Teoremine göre $\left(\frac{\chi_f^2 - f}{\sqrt{2f}}\right)$ 'in dağılımı limit durumunda

$N(0,1)$ 'dir. Buna göre

$$G_f((1+r)f) - G_f((1-r)f) \geq 1 - \alpha$$

eşitliğinden

$$\Phi\left(r\sqrt{\frac{f}{2}}\right) - \Phi\left(-r\sqrt{\frac{f}{2}}\right) \geq 1 - \alpha$$

elde edilir. f serbestlik derecesi bu eşitlikten daha kolay elde edilir.

Teorem 2.3: $P\left(\left|\frac{\hat{\sigma}^2 - \sigma^2}{\sigma^2}\right| \leq r\right) \geq 1 - \alpha$ koşulunu sağlayan örneklem genişliği n ; μ

anakütle parametresinin bilinip bilinmemesine bağlı olarak $[f^*] + 1$ veya $[f^*] + 2$ 'dir.

Burada $f^* = \frac{2Z_{\alpha/2}^2}{r^2}$ 'dir.

İspat:

$$G_f((1+r)f) - G_f((1-r)f) \geq 1 - \alpha$$

$$\Phi\left(\frac{(1+r)f - f}{\sqrt{2f}}\right) - \Phi\left(\frac{(1-r)f - f}{\sqrt{2f}}\right) \geq 1 - \alpha$$

$$\Phi\left(\frac{f(1+r-1)}{\sqrt{2f}}\right) - \Phi\left(\frac{f(1-r-1)}{\sqrt{2f}}\right) \geq 1 - \alpha$$

$$\Phi\left(\frac{\sqrt{f^2} \cdot r}{\sqrt{2f}}\right) - \Phi\left(\frac{-r\sqrt{f^2}}{\sqrt{2f}}\right) \geq 1 - \alpha$$

$$\Phi\left(r\sqrt{\frac{f}{2}}\right) - \Phi\left(-r\sqrt{\frac{f}{2}}\right) \geq 1 - \alpha$$

$$\Phi\left(r\sqrt{\frac{f}{2}}\right) - \left[1 - \Phi\left(r\sqrt{\frac{f}{2}}\right)\right] \geq 1 - \alpha$$

$$2\Phi\left(r\sqrt{\frac{f}{2}}\right) \geq 2 - \alpha$$

$$\Phi\left(r\sqrt{\frac{f}{2}}\right) \geq 1 - \frac{\alpha}{2} \rightarrow r\sqrt{\frac{f}{2}} = Z_{\alpha/2}$$

$$f^* = \frac{2Z_{\alpha/2}^2}{r^2}$$

σ^2 bilindiğinde μ hakkında hipotez testleri

α önemlilik düzeyinde $H_a : \mu > \mu_0$ tek yanlı alternatife karşı $H_0 : \mu = \mu_0$ sıfır

hipotezi test edilmek istensin. Kritik alan; $\bar{x} > \mu_0 + Z_\alpha \left(\frac{\sigma}{\sqrt{n}} \right)$ olarak biliniyor.

Alternatif $\mu = \mu_1 (> \mu_0)$ 'de testin belirli $(1 - \beta)$ güce sahip olması için örneklem genişliği n aşağıdaki teoremden belirtildiği gibi bulunur.

Teorem 2.4: α önemlilik düzeyinde $H_0 : \mu = \mu_0$ hipotezi $\mu = \mu_1 (> \mu_0)$ tek yanlı alternatifine karşı test edilirken $(1 - \beta)$ güç elde etmek için gereken örneklem genişliği

$$n; \lceil n^* \rceil + 1 \text{ 'e eşittir ve } n^* = \left[\frac{\sigma}{(\mu_1 - \mu_0)} (Z_\alpha + Z_\beta) \right]^2 \text{ 'dır.}$$

İspat:

$H_a : \mu = \mu_1 (> \mu_0)$ alternatifine karşı test edilen $H_0 : \mu = \mu_0$ sıfır hipotezi için kritik alan Neyman Pearson teoremine (Freund, 1999) göre

$$\bar{x} > \mu_0 + Z_\alpha \left(\frac{\sigma}{\sqrt{n}} \right)$$

bulunur. Buna göre

$$P \left(\bar{x} > \mu_0 + Z_\alpha \left(\frac{\sigma}{\sqrt{n}} \right) \mid \mu = \mu_1 \right) = 1 - \beta$$

$$P \left(\frac{\mu_1 - \mu_0}{\sigma / \sqrt{n}} > Z_\alpha \right) = 1 - \beta$$

$$\Phi \left(-Z_\alpha - \frac{(\mu_1 - \mu_0)}{\sigma / \sqrt{n}} \right) = 1 - \beta$$

$$\frac{(\mu_1 - \mu_0)}{\sigma / \sqrt{n}} = Z_\alpha + Z_\beta$$

$$n^* = \left[\frac{\sigma}{(\mu_1 - \mu_0)} (Z_\alpha + Z_\beta) \right]^2$$

Teoremden verilen örneklem genişliği formülü, $H_a : \mu < \mu_0$ tek yanlı alternatife karşı $H_0 : \mu = \mu_0$ hipotezinin testinde de kullanılabilir.

α önemlilik düzeyinde $H_a : \mu \neq \mu_0$ iki yanlı alternatifine karşı $H_0 : \mu = \mu_0$ hipotezi test edilmek istensin. Bu durumda kritik alan; $\left| \bar{x} - \mu_0 \right| > Z_{\alpha/2} \left(\frac{\sigma}{\sqrt{n}} \right)$ 'dir.

$\mu = \mu_1 (\neq \mu_0)$ alternatifinde belirli bir $(1 - \beta)$ güç elde etmek için aşağıda verilen koşulu sağlayan örneklem genişliği belirlenir.

$$P \left(\left| \bar{x} - \mu_0 \right| > Z_{\alpha/2} \left(\frac{\sigma}{\sqrt{n}} \right) \mid \mu = \mu_1 \right) = 1 - \beta \quad (2.4)$$

Teorem 2.5: α önemlilik düzeyinde $H_0 : \mu = \mu_0$ sıfır hipotezinin iki yanlı testi için $\mu = \mu_1 (\neq \mu_0)$ alternatifinde $(1 - \beta)$ güç elde etmek için (2.4) koşulunun sağlanmasında

gereken örneklem genişliği $[n^*] + 1$ 'e eşittir ve $n^* = \left[\frac{(Z_{\alpha/2} + Z_\beta) \sigma}{|\mu_1 - \mu_0|} \right]^2$ eşitliği ile

bulunur.

İspat:

$$P \left(\left| \bar{x} - \mu_0 \right| > Z_{\alpha/2} \left(\frac{\sigma}{\sqrt{n}} \right) \mid \mu = \mu_1 \right) = 1 - \beta \text{ koşulu}$$

$$\Phi \left(\underbrace{-Z_{\alpha/2} - \frac{(\mu_1 - \mu_0)}{\sigma / \sqrt{n}}}_X \right) + 1 - \Phi \left(\underbrace{Z_{\alpha/2} - \frac{(\mu_1 - \mu_0)}{\sigma / \sqrt{n}}}_Y \right) = 1 - \beta \quad (2.5)$$

şeklinde yazılabilir.

$\mu_1 > \mu_0$ ise X; sıfıra yaklaşır. Buna göre (2.5) eşitliği aşağıdaki gibi olur:

$$1 - \Phi \left(Z_{\alpha/2} - \frac{\mu_1 - \mu_0}{\sigma / \sqrt{n}} \right) = 1 - \beta$$

$$Z_{\alpha/2} - \frac{\mu_1 - \mu_0}{\sigma / \sqrt{n}} = -Z_\beta$$

$\mu_1 < \mu_0$ ise Y; 1'e yaklaşır. Buna göre (2.5) eşitliği aşağıdaki gibi olur:

$$-Z_{\alpha/2} - \frac{\mu_1 - \mu_0}{\sigma / \sqrt{n}} = Z_\beta$$

İki eşitlik birleştirilirse;

$$Z_{\alpha/2} + Z_{\beta} = \frac{|\mu_1 - \mu_0|}{\sigma / \sqrt{n}} \text{ olur.}$$

Buradan örneklem genişliği aşağıdaki eşitlikle hesaplanır:

$$n^* = \left[\frac{(Z_{\alpha/2} + Z_{\beta})\sigma}{|\mu_1 - \mu_0|} \right]^2$$

σ^2 bilinmediğinde μ hakkında hipotez testleri

α önemlilik düzeyinde $H_a : \mu > \mu_0$ tek yanlı alternatife karşı $H_0 : \mu = \mu_0$ sıfır

hipotezinin testi için kritik alan; $\bar{x} > \mu_0 + t_{\alpha, (n-1)} \cdot \frac{s}{\sqrt{n}}$ 'dir. $\mu = \mu_1 (> \mu_0)$ 'de $(1 - \beta)$ güç istenmesi durumunda,

$$P\left(\bar{x} > \mu_0 + t_{\alpha, (n-1)} \cdot \frac{s}{\sqrt{n}} \mid \mu = \mu_1\right) = 1 - \beta$$

koşulu elde edilir. Bu eşitliğin

$$P\left(T_{(n-1)} > t_{\alpha, (n-1)} - \sqrt{n} \frac{(\mu_1 - \mu_0)}{\sigma}\right) = 1 - \beta$$

şeklinde yazılması S ile σ tahmini yapmayı basitleştirebilir. Bu koşuldan

$$t_{\alpha, (n-1)} - \sqrt{n} \frac{(\mu_1 - \mu_0)}{\sigma} = -t_{\beta, (n-1)}$$

eşitliği elde edilir. Bilinmeyen σ için üst sınırın σ_u olduğu biliniyorsa,

$$t_{\alpha, (n-1)} - \sqrt{n} \frac{(\mu_1 - \mu_0)}{\sigma} = -t_{\beta, (n-1)}$$

eşitliği tekrarlanarak n için çözülebilir. Bu çözümü veren aşağıdaki eşitlik en küçük pozitif tamsayıyı verir:

$$n = \left[\frac{\sigma_u (t_{\alpha, (n-1)} + t_{\beta, (n-1)})}{(\mu_1 - \mu_0)} \right]^2 \quad (2.6)$$

Bilinmeyen σ için σ_u üst sınırının bilinmesi mümkün değildir. Üstelik σ_u , uç değer olarak belirlenirse örneklem genişliği n çok büyük olur. Böylece örneklem genişliği faktörünün amacı anlamsız olur.

S ile σ tahmin edilmesi, sadece büyük n için geçerli olan (2.6) eşitliğinin elde edilmesinde daha karmaşık durumlara neden olmaktadır. Bütün bu etkenlerin hepsi, σ^2 bilinmediğinde μ hakkında hipotez testi yapma durumunda örneklem genişliğinin belirlenmesi sorununa uygun cevap bulmak için (2.6) eşitliğinin kullanılamamasına neden olur.

Örneklem genişliğini belirlemede ikinci yöntem olarak iki aşamalı t testi kullanılabilir:

1. İlk olarak n_1 genişliğinde s_1^2 varyanslı bir örneklem alınır.

$$2. \quad c = \left[\frac{\mu_1 - \mu_0}{t_{\alpha, (n-1)} + t_{\beta, (n-1)}} \right]^2$$

$$3. \quad n = \max \left\{ n_1, \left[\frac{s_1^2}{c} \right] + 1 \right\}$$

4. Gerekirse $(n - n_1)$ kadar daha gözlem alınır. Bu durumda toplam örneklem genişliği n 'in ortalaması $\bar{x}$ olur.

5. α önemlilik düzeyinde $H_a : \mu > \mu_0$ tek yanlı alternatifine karşı $H_0 : \mu = \mu_0$ sıfır hipotezinin testi için en az $(1 - \beta)$ güç gerekirse kritik alan;

$$\bar{x} > \mu_0 + t_{\alpha, (n_1-1)} \cdot \frac{s}{\sqrt{n}} \text{ olur.}$$

Benzer yaklaşım $H_a : \mu < \mu_0$ tek yanlı alternatifine karşı $H_0 : \mu = \mu_0$ sıfır hipotezinin testinde de geçerlidir.

$H_a : \mu \neq \mu_0$ iki yanlı alternatifine karşı $H_0 : \mu = \mu_0$ hipotezinin testi için iki aşamalı yöntem aşağıdaki gibi olur:

1. İlk olarak n_1 genişliğinde s_1^2 varyanslı bir örneklem alınır.

$$2. \quad c = \left[\frac{|\mu_1 - \mu_0|}{t_{\alpha, (n-1)} + t_{\beta, (n-1)}} \right]^2$$

3. $n = \max \left\{ n_1, \left\lceil \frac{s_1^2}{c} \right\rceil + 1 \right\}$
4. Gerekirse $(n - n_1)$ kadar daha gözlem alınır. Bu durumda toplam örneklem genişliği n 'in ortalaması $\bar{x}$ olur.
5. α önemlilik düzeyinde $H_a : \mu > \mu_0$ tek yanlı alternatifine karşı $H_0 : \mu = \mu_0$ sıfır hipotezinin testi için en az $(1 - \beta)$ güç gerekirse kritik alan; $\left| \bar{x} - \mu_0 \right| > t_{\alpha/2, (n-1)} \cdot \frac{s_1}{\sqrt{n}}$ olur. (Desu & Raghavarao, 1990)

σ^2 hakkında hipotez testleri

Varyans testinde genellikle $H_a : \sigma^2 > \sigma_0^2$ tek yanlı alternatifine karşı $H_0 : \sigma^2 = \sigma_0^2$ sıfır hipotezi test edilir. α önemlilik düzeyinde kritik alan; $\hat{\sigma}^2 > \frac{\sigma_0^2 \chi_{(1-\alpha), f}^2}{f}$ dir. $\chi_{(1-\alpha), f}^2$; χ^2 tablo değeridir. $\sigma^2 = \sigma_1^2 (> \sigma_0^2)$ alternatifinde testin $(1 - \beta)$ güce sahip olması için

$$P \left(\hat{\sigma}^2 > \frac{\hat{\sigma}_0^2 \chi_{(1-\alpha), f}^2}{f} \mid \sigma^2 = \sigma_1^2 \right) = 1 - \beta$$

koşulunu sağlayacak n seçilir. Serbestlik derecesi f 'i bulmak için

$$\sigma_0^2 \chi_{(1-\alpha), f}^2 = \sigma_1^2 \chi_{\beta, f}^2 \quad (2.7)$$

ilişkisi çözülür.

Teorem 2.6: α önemlilik düzeyinde $H_a : \sigma^2 = \sigma_1^2 (> \sigma_0^2)$ tek yanlı alternatifine karşı $H_0 : \sigma^2 = \sigma_0^2$ sıfır hipotezinin testinde $(1 - \beta)$ güç elde etmek için gereken n örneklem genişliği, μ ortalamasının bilinip bilinmemesine bağlı olarak f veya $(f + 1)$ 'e eşittir. f ; (2.7) eşitliğinden bulunan pozitif en küçük tamsayıdır.

İspat:

$$P\left(\hat{\sigma}^2 > \frac{\hat{\sigma}_0^2 \chi_{(1-\alpha),f}^2}{f} \mid \sigma^2 = \sigma_1^2\right) = 1 - \beta$$

$$P\left(\frac{f\hat{\sigma}^2}{\sigma^2} > \frac{\hat{\sigma}_0^2 \chi_{(1-\alpha),f}^2}{\sigma^2} \mid \sigma^2 = \sigma_1^2\right) = 1 - \beta$$

$$\frac{\hat{\sigma}_0^2 \chi_{(1-\alpha),f}^2}{\sigma_1^2} = \chi_{\beta,f}^2$$

$$\hat{\sigma}_0^2 \chi_{(1-\alpha),f}^2 = \sigma_1^2 \chi_{\beta,f}^2$$

χ^2 değeri; $\chi_{\beta,f}^2 = \left(\Phi^{-1}(p) + \sqrt{2f-1}\right)^2 \cdot \frac{1}{2}$ şeklinde normal değerlerle tahmin

edilebilir.

(2.7) eşitliğinden $f^* = \frac{1}{2} \cdot \left(\frac{\lambda Z_\alpha + Z_\beta}{\lambda - 1}\right)^2 + \frac{1}{2}$ bulunur. Burada λ^2 , $\lambda = \frac{\sigma_0^2}{\sigma_1^2}$ 'ye

eşittir.

Teorem 2.7: α önemlilik düzeyinde $H_a : \sigma^2 = \sigma_1^2 (> \sigma_0^2)$ tek yanlı alternatifine karşı $H_0 : \sigma^2 = \sigma_0^2$ sıfır hipotezinin testinde $(1 - \beta)$ güç elde etmek için gereken örneklem genişliği, μ 'nün bilinip bilinmemesine bağlı olarak $[f^*] + 1$ veya $[f^*] + 2$ 'ye eşittir.

Burada $f^* = \frac{1}{2} \cdot \left(\frac{\lambda Z_\alpha + Z_\beta}{\lambda - 1}\right)^2 + \frac{1}{2}$ 'dır.

İspat:

$\chi_{\beta,f}^2 = \frac{1}{2} \cdot \left(\Phi^{-1}(p) + \sqrt{2f-1}\right)^2$ ise (2.7) eşitliği şu şekilde yazılabilir:

$$\frac{1}{2} \cdot \sigma_0^2 \left[\Phi^{-1}(1-\alpha) + \sqrt{2f-1} \right] = \frac{1}{2} \cdot \sigma_1^2 \left[\Phi^{-1}(\beta) + \sqrt{2f-1} \right]$$

$$\sigma_0^2 (Z_\alpha + \sqrt{2f-1}) = \sigma_1^2 (-Z_\beta + \sqrt{2f-1})$$

$$\sigma_0^2 Z_\alpha + \sigma_0^2 \sqrt{2f-1} = -\sigma_1^2 Z_\beta + \sigma_1^2 \sqrt{2f-1}$$

$$\sigma_0^2 Z_\alpha + \sigma_1^2 Z_\beta = \sqrt{2f-1} (\sigma_1^2 - \sigma_0^2)$$

$$\lambda = \frac{\sigma_0^2}{\sigma_1^2} \text{ olsun.}$$

$$\lambda \sigma_1^2 Z_\alpha + \sigma_1^2 Z_\beta = \sqrt{2f-1} (\sigma_1^2 - \lambda \sigma_1^2)$$

$$\sigma_1^2 (\lambda Z_\alpha + Z_\beta) = -\sqrt{2f-1} \sigma_1^2 (\lambda - 1)$$

$$f^* = \frac{1}{2} \cdot \left(\frac{\lambda Z_\alpha + Z_\beta}{\lambda - 1} \right)^2 + \frac{1}{2}$$

2.1.2. Bernoulli Dağılımlı Bir Anakütleden Örnek Seçimi

X kesikli rastlantı değişkeni

$$f(x, \theta) = \theta^x (1 - \theta)^{1-x}, x = 0, 1$$

olasılık fonksiyonuna sahip olduğunu varsayalım. Bernoulli dağılımının ortalaması θ , varyansı $\theta(1 - \theta)$ 'dır. $X_1, X_2, \dots, X_n$ rastlantı değişkenlerinden oluşan rastsal örneklem olsun. $Y = \sum_{i=1}^n X_i$, n ve θ parametrelili Binom dağılımına sahiptir. Binom dağılımının olasılık fonksiyonu aşağıda verilmiştir:

$$f(y; n, \theta) = \binom{n}{y} \theta^y (1 - \theta)^{n-y}, y = 0, 1, \dots, n$$

Buna göre Y rastlantı değişkeninin ortalaması; $n\theta$, varyansı; $n\theta(1 - \theta)$ 'dır.

θ 'nın yansız tahmin edicisi $\frac{Y}{n}$ 'e eşittir.

θ 'nın tahmini

θ anakütle parametresini $\hat{\theta} = \frac{Y}{n}$ ile tahmin etmek için gereken örneklem genişliğini belirlerken mutlak hata kontrol edilir. d ve α pozitif sabitler olmak üzere

$$P\left(\left|\frac{Y}{n} - \theta\right| \leq d\right) \geq 1 - \alpha \quad (2.8)$$

koşulunu sağlayan n seçilir. (2.8) eşitsizliğinin sol tarafı aşağıdaki gibi yazılabilir:

$$P\left(\left[n(\theta-d)\right]+1 \leq Y \leq \left[n(\theta+d)\right]\right) \equiv P(y_1 \leq Y \leq y_2) \\ = \sum_{y=y_1}^{y_2} \binom{n}{y} \theta^y (1-\theta)^{n-y}$$

Bilinmeyen θ parametresine bağlı bu olasılık $\theta = 0.5$ olduğunda en küçük değeri alır. Olasılığın $(1-\alpha)$ 'dan küçük olduğu durumda uygulanabilen bu eşitlikte $\theta = 0.5$ alındığında yeterli örneklem genişliği bulunmuş olur.

Merkezi Limit Teoremine göre;

$$\frac{\left|\frac{Y}{n} - \theta\right|}{\sqrt{\theta(1-\theta)/n}}$$

değişkeni limit durumunda (asimptotik olarak) Standart Normal Dağılıma sahiptir. Buradan

$$\Phi\left(\frac{d}{\sqrt{\theta(1-\theta)/n}}\right) - \Phi\left(\frac{-d}{\sqrt{\theta(1-\theta)/n}}\right) \geq 1-\alpha \\ \Phi\left(\frac{d}{\sqrt{\theta(1-\theta)/n}}\right) \geq 1-\frac{\alpha}{2}$$

eşitliği elde edilir.

Bu koşulu sağlayan örneklem genişliği,

$$n \geq \frac{\theta(1-\theta)Z_{\alpha/2}^2}{d^2}$$

koşulunu da sağlamalıdır.

Anakütle parametresi θ bilinmediğinde ve $\theta = 0.5$ alındığında $\theta(1-\theta)$ en büyük olacağından örneklem genişliği n 'in yeterli tahmini olarak $\left[n^*\right]+1$ kullanılabilir.

Burada $n^* = \frac{Z_{\alpha/2}^2}{4d^2}$ 'dir.

Teorem 2.8: θ bilinmeyen anakütle parametresinin tahmini için gereken örneklem genişliği, Bernoulli dağılımının başarı olasılığının $\hat{\theta} = \frac{Y}{n}$ alınması ve

$P\left(\left|\frac{Y}{n}-\theta\right|\leq d\right)\geq 1-\alpha$ koşulunun sağlanmasıyla $[n^*]+1$ olarak bulunur. n^* ,

yukarıdaki gibidir.

İspat:

$$P\left(\left|\frac{Y}{n}-\theta\right|\leq d\right)\geq 1-\alpha$$

$$P\left(-d\leq \frac{Y}{n}-\theta\leq d\right)\geq 1-\alpha$$

$$P\left(-\frac{d}{\sqrt{\theta(1-\theta)/n}}\leq \frac{\frac{Y}{n}-\theta}{\sqrt{\theta(1-\theta)/n}}\leq \frac{d}{\sqrt{\theta(1-\theta)/n}}\right)\geq 1-\alpha$$

$$\Phi\left(\frac{d}{\sqrt{\theta(1-\theta)/n}}\right)-\Phi\left(-\frac{d}{\sqrt{\theta(1-\theta)/n}}\right)\geq 1-\alpha$$

$$\Phi\left(\frac{d}{\sqrt{\theta(1-\theta)/n}}\right)-1+\Phi\left(\frac{d}{\sqrt{\theta(1-\theta)/n}}\right)\geq 1-\alpha$$

$$\Phi\left(\frac{d}{\sqrt{\theta(1-\theta)/n}}\right)\geq 1-\frac{\alpha}{2}$$

θ bilinmiyor ancak $\theta = 0.5$ alınırsa;

$$\frac{d}{\sqrt{(0.5)(0.5)/n}}=Z_{\alpha/2}$$

$$n^*=\frac{Z_{\alpha/2}^2}{4d^2}$$

θ hakkında hipotez testi

α önemlilik düzeyinde $H_a:\theta>\theta_0$ tek yanlı alternatifine karşı $H_0:\theta=\theta_0$

hipotezi test edilmek istensin. Kritik alan; $Y>c$ 'dir. c ;

$$P(Y>c\mid\theta=\theta_0)\leq\alpha$$

veya

$$I_{\theta_0}(c+1,n-c)\leq\alpha$$

eşitliğinden belirlenir.

$I_x(a, b)$ tamamlanmamış Beta fonksiyonudur.

$$I_x(a, b) = \frac{\int_0^x u^{a-1} (1-u)^{b-1} du}{\int_0^1 u^{a-1} (1-u)^{b-1} du}$$

$H_a : \theta = \theta_1 (> \theta_0)$ alternatifinde $(1 - \beta)$ güç istenmesiyle

$$P(Y > c \mid \theta = \theta_1) \geq 1 - \beta$$

koşulu elde edilir. Bu koşuldan örneklem genişliği bulunur. n ve c aşağıdaki iki koşulu sağlamalıdır:

$$I_{\theta_0}(c+1, n-c) \leq \alpha \quad I_{\theta_1}(c+1, n-c) \geq 1 - \beta$$

$\frac{Y}{n}$ 'e arcsin dönüşümü yapılarak n hesaplanır. Standart Normal Dağılım terimlerinde,

$$P(Y > c \mid \theta = \theta_0) \leq \alpha, P(Y > c \mid \theta = \theta_1) \geq 1 - \beta$$

olasılıkları kullanılır.

Bilindiği gibi

$$Z = 2\sqrt{n} \left(\arcsin \sqrt{\hat{\theta}} - \arcsin \sqrt{\theta} \right)$$

değişkeni asimptotik olarak Standart Normal Dağılımlıdır. $\text{Arcsin}(\cdot)$ fonksiyonu

$\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ aralığında tanımlıdır. Bu nedenle n örneklem genişliği;

$$1 - \Phi \left\{ 2\sqrt{n} \left(\arcsin \sqrt{c/n} - \arcsin \sqrt{\theta_0} \right) \right\} = \alpha$$

ve

$$1 - \Phi \left\{ 2\sqrt{n} \left(\arcsin \sqrt{c/n} - \arcsin \sqrt{\theta_1} \right) \right\} = 1 - \beta$$

eşitliklerini sağlar. Bu eşitliklerden n örneklem genişliği;

$$n^* = \left\{ \frac{(Z_\alpha + Z_\beta)}{2(\arcsin \sqrt{\theta_1} - \arcsin \sqrt{\theta_0})} \right\}^2$$

olarak bulunur. Gereken örneklem genişliği, $[n^*] + 1$ 'dir.

2.1.3. Tahmin İçin Bayes Yaklaşımı

Bu bölüme kadarki tartışılan tüm bölümlerde klasik yöntemler kullanıldı. Olasılıklar, nisbi frekanslar olarak yorumlandı. Araştırmacılar, parametre değeri hakkında önsel bilgilere sahip olduklarında, örneklerle sağlanan bilgilerin ışığında bu önsel bilgileri değiştirmek isteyebilirler. Bu yolla çıkarım yapmada olasılıklar, genellikle kişisel bilgiler olarak yorumlanır. Parametre tahmini bağlamında örneklem genişliği tahmini için, aşağıdaki Bayesci yaklaşım da bir yöntemdir.

Burada problemi kesin ve açık olarak belirtmek için Bayes yaklaşımı anahatlarıyla verilecektir. Bu yaklaşım Guttman v.d.(1982)'nin görüşleri doğrultusundadır.

X rastlantı değişkenlerinden oluşan n genişlikli rastsal örneklemde, olasılık veya olasılık yoğunluk fonksiyonu $f(x/\underline{\theta})$ ile, $\underline{\theta}$ 'dan sağlanan bilgilerle gözlenecektir. $\underline{\theta}$ parametresi, Θ kümesinde değerler alan sürekli değişkenlerdir. Önsel bilgiler; $P(\underline{\theta})$ ile tanımlı bir önsel dağılım şeklinde belirtilebilir. X rastlantı değişkenlerinden oluşan rastsal örneklem $\underline{X}$ olsun. Verilen $\underline{X}$, $\underline{\theta}$ 'nın sonsal olasılık yoğunluk fonksiyonu aşağıdaki gibidir:

$$\pi^*(\underline{\theta} \setminus \underline{X}) = \frac{P(\underline{\theta})L(\underline{\theta} \setminus \underline{X})}{\int_{\Theta} P(\underline{\theta})L(\underline{\theta} \setminus \underline{X})d\underline{\theta}}$$

L ; örneklem olabirlik fonksiyonudur. Sonsal dağılım, toplanan tüm bilgilerle, $\underline{\theta}$ değeri ile ilgilidir.

θ 'nın θ_1 bileşeni toplansın. θ_1 'in marjinal sonsal yoğunluğu $\pi(\theta_1 \setminus X)$ 'tir. $\pi(\theta_1 \setminus X)$, çıkarım amacı için kullanılacaktır. Bu sonsal dağılımın ortalaması, nokta tahmini gibi kullanılır ve θ_1 'in Bayes tahmini olarak adlandırılır. Araştırmacının amacı nokta tahmin ise; örneklem genişliği tanımını aşağıdaki gibi açıkça belirtilebilir:

P önseline ve $\underline{X}$ rastlantı örneğine dayalı θ_1 'in sonsal ortalaması $\hat{\theta}_1(\underline{x}; p)$ olsun. d ve α pozitif sabitler olmak üzere $P\left(\left|\theta_1 - \hat{\theta}_1(\underline{x}; p)\right| \leq d \setminus \underline{x}\right) \geq 1 - \alpha$ koşulunu sağlayan örneklem genişliği seçilir. $\hat{\theta}_1(\underline{x}; p) \mp d$ koşulunu sağlamakla benzer olan bu yaklaşımda örneklem genişliği seçimi, θ_1 için $(1 - \alpha)$ düzeyinin sonsal aralığıdır. Diğer bir deyişle, belirli $(1 - \alpha)$ sonsal aralık olan Bayes tahmini etrafında, belirli $2d$ genişlikli aralık sağlanmak istenir.

Alternatif olarak, seçilen pozitif $r(<1)$ ve α için, $P\left(\left|\frac{\theta_1 - \hat{\theta}_1(\underline{x}; p)}{\hat{\theta}_1(\underline{x}; p)}\right| \leq r \setminus \underline{x}\right) \geq 1 - \alpha$ koşulunu sağlayacak örneklem genişliği seçilir.

Örneklem genişliği bu şekilde seçildiği zaman, $\left([1 - r]\hat{\theta}_1(\underline{x}; p), [1 + r]\hat{\theta}_1(\underline{x}; p)\right)$ aralığı,

θ_1 için $(1 - \alpha)$ düzeyli sonsal aralıktır. Bu aralık; $\frac{\text{Üst nokta}}{\text{Alt nokta}} - 1 = 1 - \frac{\text{Alt nokta}}{\text{Üst nokta}} = r$

özelliğine sahiptir.

2.2. İki Örneklem Durumunda Örneklem Genişliği Belirleme

X ve Y , $F(x)$ ve $G(y)$ dağılım fonksiyonlarına sahip bağımsız rastlantı değişkenleri olsun. Bu rastlantı değişkenlerinin parametreleri veya dağılımları, X ve Y 'ye dayanarak çekilen iki örnekleme göre karşılaştırılabilir. Örneğin, $E(x) - E(y)$ ortalama farkları veya varyanslarının oranlarından çıkarımlar elde edilebilir.

Bu bölümde, belirli koşulları sağlayacak çıkarımlar elde etmek için gereken örneklem genişliklerini tanımlama problemleri incelenecektir.

2.2.1. Normal Dağılımlı İki Anakütleden Örnek Seçimi

X ; μ_1 ortalamalı ve σ_1^2 varyanslı normal dağılıma, Y de μ_2 ortalamalı ve σ_2^2 varyanslı normal dağılıma sahip olduğunu varsayalım. $X_1, X_2, \dots, X_{n_1}$; ortalaması μ_1 olan normal dağılımlı bir anakütleden çekilen n_1 genişlikli bir rastsal örneklem, $Y_1, Y_2, \dots, Y_{n_2}$; ortalaması μ_2 olan normal dağılımlı bir anakütleden çekilen n_2 genişlikli rastsal örneklem ise; bu iki örneklemin ortalama ve varyansları aşağıdaki gibidir:

$$\begin{aligned}\bar{X} &= \sum_{i=1}^{n_1} \frac{X_i}{n_1} & S_1^2 &= \sum_{i=1}^{n_1} \frac{(X_i - \bar{X})^2}{n_1 - 1} \\ \bar{Y} &= \sum_{i=1}^{n_2} \frac{Y_i}{n_2} & S_2^2 &= \sum_{i=1}^{n_2} \frac{(Y_i - \bar{Y})^2}{n_2 - 1}\end{aligned}$$

Bu iki rastsal örneklemin ortalama farkları ve varyans oranlarına dayalı hipotez testleri ve tahminler aşağıdaki bölümlerde incelenmiştir:

σ_1^2 ve σ_2^2 Bilindiğinde $\mu_1 - \mu_2$ Tahmini

Ortalamalar farkı $\mu_1 - \mu_2$ 'nin yansız tahmin edicisi; $(\bar{X} - \bar{Y})$ 'dir. $(\mu_1 - \mu_2)$ ortalama, $\left(\frac{\sigma_1^2}{n_1}, \frac{\sigma_2^2}{n_2}\right)$ varyanslı normal dağılıma sahiptir. α önemlilik düzeyinde, d ve α sabitler olmak üzere;

$$P\left(\left|(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)\right| \leq d\right) \geq 1 - \alpha \quad (2.9)$$

koşulunu sağlayan n_1 ve n_2 tanımlanır. Burada ikinci bir koşula gereksinim duyulur. Maliyet ve diğer faktörler dikkate alındığında, uygun k için genellikle, $n_1 = kn_2$ koşulu kabul edilir. Ama pek çok problemde, $k=1$ özel durumu (eşit örneklem genişliği) alınır. n_1 ve n_2 hakkında aşağıdaki sonuçlar ortaya çıkar:

Teorem 2.9: $P\left(\left|(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)\right| \leq d\right) \geq 1 - \alpha$ koşulunu sağlayan iki örneklem

genişliği; $n_1 = \lceil kn^* \rceil + 1$ ve $n_2 = \lceil n^* \rceil + 1$ 'dir ve $n^* = \frac{\left\{\left(\sigma_1^2 / k\right) + \sigma_2^2\right\} Z_{\alpha/2}^2}{d^2}$ 'dir.

İspat:

$$P\left(\left|(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)\right| \leq d\right) \geq 1 - \alpha \quad n_1 = kn_2$$

$$P\left(-d \leq (\bar{X} - \bar{Y}) - (\mu_1 - \mu_2) \leq d\right) \geq 1 - \alpha$$

$$P\left(-\frac{d}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \leq Z \leq \frac{d}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}\right) \geq 1 - \alpha$$

$$\Phi\left(\frac{d}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}\right) - \Phi\left(-\frac{d}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}\right) \geq 1 - \alpha$$

$$2\Phi\left(\frac{d}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}\right) \geq 1 - \alpha$$

$$\frac{d}{\sqrt{\frac{\sigma_1^2}{kn} + \frac{\sigma_2^2}{n}}} = Z_{\alpha/2}$$

$$n^* = \frac{\left\{(\sigma_1^2/k) + \sigma_2^2\right\} Z_{\alpha/2}^2}{d^2}$$

$\sigma_1^2 = \sigma_2^2 (= \sigma^2)$ olduğunda, maliyet değişimine bakılmadan $n_1 = n_2$ alınır. Bu

durumda; $n_1 = n_2 = \left\lceil \frac{2\sigma^2 Z_{\alpha/2}^2}{d^2} \right\rceil + 1$ olur.

İspat:

$$\frac{d}{\sqrt{\frac{\sigma}{n} + \frac{\sigma}{n}}} = Z_{\alpha/2}$$

$$n^* = \frac{2\sigma^2 Z_{\alpha/2}^2}{d^2}$$

$\sigma_1^2 = \sigma_2^2$ Olduğunda ve Ortak Varyans Bilinmediğinde $\mu_1 - \mu_2$ Tahmini

Bu durumda $P\left(\left|(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)\right| \leq d\right) \geq 1 - \alpha$ koşulunu sağlamak için İki Aşamalı yöntem(1945) kullanmak gerekir.

1. İlk olarak her iki anakütleden $n_1 (\geq 2)$ genişlikli örneklem alınsın.
2. s_p^2 , bu iki örneklem birleştirilmiş varyansı olsun.
3. $c = \frac{1}{2} \left\{ \frac{d}{t_{\alpha/2, (2n_1-2)}} \right\}^2$ hesaplanır.
4. $n = \max \left\{ n_1, \left\lceil \frac{s_p^2}{c} \right\rceil + 1 \right\}$ olur.
5. Gerekliyse her iki anakütleden $(n - n_1)$ kadar daha gözlem alınır. $\bar{X}$ ve $\bar{Y}$, n gözlemlerin tümüne dayanan iki örneklem ortalamasıdır.

Bu durumda $(\mu_1 - \mu_2)$ 'nin tahmin edicisi olarak $(\bar{X} - \bar{Y})$, $P\left(\left|(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)\right| \leq d\right) \geq 1 - \alpha$ koşulunu sağlar.

$\sigma_1^2 \neq \sigma_2^2$ Olduğunda ve Varyanslar Bilinmediğinde $(\mu_1 - \mu_2)$ Tahmini

Koopmans ve Qualls(1971), bu durum için, $P\left(\left|(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)\right| \leq d\right) \geq 1 - \alpha$ koşulunu sağlayan aşağıda tanımlanmış İki Aşamalı yöntemi ileri sürmüştür:

1. İlk olarak $n_{11} (\geq 2)$ ve $n_{21} (\geq 2)$ genişlikli bağımsız örneklem alınsın. s_{11}^2 ve s_{21}^2 , bu örneklerin varyansı olsun.
2. σ_{1u}^2 ve σ_{2u}^2 olan σ_1^2 ve σ_2^2 'nin $\%100\sqrt{1 - (\alpha/2)}$ üst güven sınırları hesaplanır.
3. n_{12} ve n_{22} ikinci aşamada iki anakütleden alınan örneklemelerin genişlikleri olmak üzere, $C = c_0 + c_1 n_{12} + c_2 n_{22}$ doğrusal maliyet fonksiyonu olsun.

$$Z_{\alpha/4} \sqrt{\frac{\sigma_{1u}^2}{n_{12}} + \frac{\sigma_{2u}^2}{n_{22}}} \leq d \text{ ile } C \text{ maliyet fonksiyonunu en küçük yapan } n_{12} \text{ ve } n_{22}$$

çözümü aşağıdaki gibidir:

$$n_{12} = \left\lceil \frac{Z_{\alpha/4}^2}{d^2} \cdot \frac{\sqrt{c_1} \sigma_{1u} + \sqrt{c_2} \sigma_{2u}}{\sqrt{c_1}} \sigma_{1u} \right\rceil + 1$$

$$n_{22} = \left\lceil \frac{Z_{\alpha/4}^2}{d^2} \cdot \frac{\sqrt{c_1} \sigma_{1u} + \sqrt{c_2} \sigma_{2u}}{\sqrt{c_2}} \sigma_{2u} \right\rceil + 1$$

4. İkinci aşama örneklemelerin ortalamaları $\bar{X}$ ve $\bar{Y}$ olsun. Bu durumda $(\mu_1 - \mu_2)$ 'nin tahmin edicisi olan $(\bar{X}_2 - \bar{Y}_2)$,

$$P\left(\left|(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)\right| \leq d\right) \geq 1 - \alpha \text{ koşulunu sağlar.}$$

Tahmin edicinin sadece bu yöntemde ikinci aşama örneklem ortalamasına bağlı olduğuna dikkat edilmelidir.

σ_1^2 / σ_2^2 Varyans Oranının Tahmini

$$\begin{aligned} \hat{\sigma}_1^2 &= \sum_{i=1}^{n_1} \frac{(X_i - \mu_1)^2}{n_1}, \mu_1 \text{ biliniyorsa} \\ &= s_1^2, \mu_1 \text{ bilinmiyorsa} \\ \hat{\sigma}_2^2 &= \sum_{j=1}^{n_2} \frac{(Y_j - \mu_2)^2}{n_2}, \mu_2 \text{ biliniyorsa} \\ &= s_2^2, \mu_2 \text{ bilinmiyorsa} \end{aligned}$$

ν_i , μ 'nün bilinip bilinmediğine bağlı olarak n_i veya $(n_i - 1)$ olsun. σ_1^2 / σ_2^2 oranının tahmin edicisi olarak genellikle $\hat{\sigma}_1^2 / \hat{\sigma}_2^2$ kullanılır ve $(\hat{\sigma}_1^2 / \hat{\sigma}_2^2) / (\sigma_1^2 / \sigma_2^2)$; ν_1 ve ν_2 serbestlik dereceli F dağılır. Bu; $(\hat{\sigma}_1^2 / \hat{\sigma}_2^2) / (\sigma_1^2 / \sigma_2^2) \square F_{\nu_1, \nu_2}$ şeklinde gösterilir. Bu durumda,

$$P\left(\left|\frac{(\hat{\sigma}_1^2 / \hat{\sigma}_2^2) - (\sigma_1^2 / \sigma_2^2)}{(\sigma_1^2 / \sigma_2^2)}\right| \leq r\right) \geq 1 - \alpha \quad (2.10)$$

koşulunun belirtilmesiyle nisbi hata kontrol edilir.

Bu koşul basitleştirilerek,

$$P(1-r \leq F_{\nu_1, \nu_2} \leq 1+r) \geq 1-\alpha$$

şeklinde yazılır. $H_{\nu_1, \nu_2}(\cdot)$, F_{ν_1, ν_2} değişkeninin dağılım fonksiyonu ise

$$P(1-r \leq F_{\nu_1, \nu_2} \leq 1+r) \geq 1-\alpha$$

koşulu,

$$H_{\nu_1, \nu_2}(1+r) - H_{\nu_1, \nu_2}(1-r) \geq 1-\alpha \quad (2.11)$$

olarak tekrar yazılabilir.

Teorem 2.10: $P\left(\left|\frac{(\hat{\sigma}_1^2/\hat{\sigma}_2^2) - (\sigma_1^2/\sigma_2^2)}{(\sigma_1^2/\sigma_2^2)}\right| \leq r\right) \geq 1-\alpha$ koşulunu sağlamak için gereken

örneklem genişliği, (2.11) eşitliğini sağlayan çözümdür. $Z = \frac{\frac{1}{2} \ln F_{\nu_1, \nu_2} - \frac{1}{2} \left(\frac{1}{\nu_2} - \frac{1}{\nu_1} \right)}{\sqrt{\frac{1}{2} \left(\frac{1}{\nu_1} + \frac{1}{\nu_2} \right)}}$

asimptotik dağılımının $N(0,1)$ olduğunun bilinmesiyle (2.11) eşitliğinin yaklaşık çözümü elde edilebilir.

$1/\nu_2$ 'nin yaklaşık olarak $1/\nu_1$ 'e eşit olduğu varsayalım. Bu durumda (2.11) eşitliği,

$$\Phi\left\{\frac{\frac{1}{2} \ln(1+r)}{\sqrt{\frac{1}{2} \left(\frac{1}{\nu_1} + \frac{1}{\nu_2} \right)}}\right\} - \Phi\left\{\frac{\frac{1}{2} \ln(1-r)}{\sqrt{\frac{1}{2} \left(\frac{1}{\nu_1} + \frac{1}{\nu_2} \right)}}\right\} \geq 1-\alpha$$

eşitsizliği ile asimptotik olarak tahmin edilebilir. Eşitsizliğin sol tarafındaki ikinci terim sonsuza gider ve eşitsizlik;

$$\Phi \left\{ \frac{\frac{1}{2} \ln(1+r)}{\sqrt{\frac{1}{2} \left(\frac{1}{\nu_1} + \frac{1}{\nu_2} \right)}} \right\} \geq 1 - \alpha$$

durumuna indirgenir. Eşitsizliğin çözümü;

$$\frac{\frac{1}{2} \ln(1+r)}{\sqrt{\frac{1}{2} \left(\frac{1}{\nu_1} + \frac{1}{\nu_2} \right)}} = Z_\alpha \text{ 'dır.}$$

Bu çözüm, $H_{\nu_1, \nu_2}(1+r) - H_{\nu_1, \nu_2}(1-r) \geq 1 - \alpha$ eşitliğinin de yaklaşık çözümüdür.

$$\frac{1}{\nu_1} + \frac{1}{\nu_2} = \frac{1}{2} \left\{ \frac{\ln(1+r)}{Z_\alpha} \right\}^2 = \tau$$

Uygun c^* için $c^* = c_1 \nu_1 + c_2 \nu_2$ yazılmasıyla çözülebilen $C = c_0 + c_1 n_{12} + c_2 n_{22}$ doğrusal maliyet fonksiyonunun en küçüklenmesiyle, ν_1 ve ν_2 için bir çözüm elde edilebilir. $(c_1 n_1 + c_2 n_2) \cdot \left(\frac{1}{\nu_1} + \frac{1}{\nu_2} \right) \geq (\sqrt{c_1} + \sqrt{c_2})^2$ Cauchy-Schwartz eşitliği kullanılarak; ν_1 ve ν_2 'nin çözümü $\nu_1^* = \frac{1}{\tau} \left(1 + \sqrt{\frac{c_2}{c_1}} \right)$, $\nu_2^* = \frac{1}{\tau} \left(1 + \sqrt{\frac{c_1}{c_2}} \right)$ olduğunda en küçük c^* elde edilir.

Teorem 2.11: $c^* = c_1 \nu_1 + c_2 \nu_2$ maliyet fonksiyonunun en küçüklenmesiyle bulunan ve

$$P \left(\left| \frac{(\hat{\sigma}_1^2 / \hat{\sigma}_2^2) - (\sigma_1^2 / \sigma_2^2)}{(\sigma_1^2 / \sigma_2^2)} \right| \leq r \right) \geq 1 - \alpha \text{ koşulunun sağlanması için gereken örneklem}$$

genişliği; μ_i 'nin bilinip bilinmemesine bağlı olarak $[\nu_i] + 1$ veya $[\nu_i] + 2$ 'dir. ν_i ;

$$\nu_1^* = \frac{1}{\tau} \left(1 + \sqrt{\frac{c_2}{c_1}} \right), \nu_2^* = \frac{1}{\tau} \left(1 + \sqrt{\frac{c_1}{c_2}} \right) \text{ ile bulunur. (i=1,2 için)}$$

$c_1 = c_2$ olduğunda yukarıdaki formüller aşağıdaki gibi olur:

$$\nu_1^* = \nu_2^* = \frac{2}{\tau} = 4 \left\{ \frac{Z_\alpha}{\ln(1+r)} \right\}^2$$

σ_1^2 ve σ_2^2 Bilindiğinde μ_1 ve μ_2 ile İlgili Hipotez Testi

$H_a: \mu_1 > \mu_2$ alternatifine karşı $H_0: \mu_1 = \mu_2$ sıfır hipotezi α düzeyinde test

edilmek istensin. Bu test için kritik alan; $\bar{X} - \bar{Y} > Z_\alpha \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$ 'dir.

$\mu_1 - \mu_2 = \delta (> 0)$ 'da, $(1 - \beta)$ güç istenirse; n_1 ve n_2 örneklem genişlikleri

$$P \left(\bar{X} - \bar{Y} > Z_\alpha \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \mid \mu_1 - \mu_2 = \delta \right) = 1 - \beta$$

koşulunu sağlamalıdır.

$$\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} = \left\{ \frac{\delta}{Z_\alpha + Z_\beta} \right\}^2 \text{ eşitliğinde tek çözüm elde etmek için } n_1 = kn_2 \text{ alınır ve}$$

$$n_2^* = \frac{\frac{\sigma_1^2}{k} + \sigma_2^2}{\left(\frac{\delta}{Z_\alpha + Z_\beta} \right)^2}, \quad n_1^* = kn_2^* \text{ bulunur.}$$

Teorem 2.12: $\mu_1 - \mu_2 = \delta (> 0)$ alternatifine karşı $H_0: \mu_1 = \mu_2$ sıfır hipotezinin α düzeyli testinde $(1 - \beta)$ güç elde etmek için gereken n_1 ve n_2 örneklem genişlikleri

$$\text{sırasıyla } \lceil kn_2^* \rceil + 1 \text{ ve } \lceil n_2^* \rceil + 1 \text{ 'e eşittir ve } n_2^* = \frac{\frac{\sigma_1^2}{k} + \sigma_2^2}{\left(\frac{\delta}{Z_\alpha + Z_\beta} \right)^2} \text{ 'dir.}$$

İspat:

$$P \left(\bar{X} - \bar{Y} > Z_\alpha \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \mid \mu_1 - \mu_2 = \delta \right) = 1 - \beta$$

$$P \left(\frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} > Z_\alpha \right) = 1 - \beta$$

$\mu_1 - \mu_2 = \delta$ ise

$$\frac{\delta}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = Z_\alpha + Z_\beta$$

$n_1 = kn_2$ ise

$$\frac{\sigma_1^2 / k}{n_2} + \frac{\sigma_2^2}{n_2} = \left(\frac{\delta}{Z_\alpha + Z_\beta} \right)^2$$

$$n_2^* = \frac{\frac{\sigma_1^2}{k} + \sigma_2^2}{\left(\frac{\delta}{Z_\alpha + Z_\beta} \right)^2}$$

Benzer sonuç bu formülde α yerine $\alpha/2$ koyularak çift yanlı test için bulunabilir.

$\sigma_1^2 = \sigma_2^2$ Olduğunda ve Ortak Varyans Bilinmediğinde μ_1 ve μ_2 ile İlgili Hipotez Testi

Her bir anakütleden n gözlem alınmış $H_a : \mu_1 > \mu_2$ tek yanlı alternatifine karşı $H_0 : \mu_1 = \mu_2$ α düzeyli testi düşünelim. Kullanılan t testinin kritik alanı;

$\bar{X} - \bar{Y} > t_{\alpha, \nu} S_p \sqrt{\frac{2}{n}}$, dir. ($\nu = 2(n-1)$, S_p^2 ; birleştirilmiş varyans)

$\mu_1 - \mu_2 = \delta$ olduğunda $(1 - \beta)$ güç istenmesiyle, n için çözülen

$$P \left(\bar{X} - \bar{Y} > t_{\alpha, \nu} S_p \sqrt{\frac{2}{n}} \mid \delta \right) = 1 - \beta$$

koşulu elde edilir. Sonra bu eşitlik;

$$P\left(T_v > t_{\alpha,v} - \frac{\delta}{S_p \sqrt{\frac{2}{n}}}\right) = 1 - \beta \quad (2.12)$$

şekline indirgenir.

S_p ile σ tahmin edilerek (2.12) eşitliğinden örneklem genişliği aşağıdaki gibi hesaplanır:

$$n = \frac{v}{2} + 1 = 2 \left(\frac{\sigma}{\delta} \right)^2 \left\{ t_{\alpha,v} + t_{\beta,v} \right\}^2$$

Teorem 2.13: $H_0 : \mu_1 = \mu_2$ sıfır hipotezi tek yanlı α düzeyli test edilirken, $\mu_1 - \mu_2 = \delta (> 0)$ olduğunda $(1 - \beta)$ güç elde etmek için gereken $n_1 = n_2 = n$ örneklem genişliği; $\left(\frac{v}{2} + 1 \right)$ 'e eşittir. v ; $n = \frac{v}{2} + 1 = 2 \left(\frac{\sigma}{\delta} \right)^2 \left\{ t_{\alpha,v} + t_{\beta,v} \right\}^2$ eşitliğini sağlayan en küçük pozitif tam sayıdır.

İspat:

$$\begin{aligned} P\left(\bar{X} - \bar{Y} > t_{\alpha,v} S_p \sqrt{\frac{2}{n}} \mid \delta\right) &= 1 - \beta \\ P\left(\frac{(\bar{X} - \bar{Y}) - \delta}{S_p \sqrt{\frac{2}{n}}} + \frac{\delta}{S_p \sqrt{\frac{2}{n}}} > t_{\alpha,v}\right) &= 1 - \beta \\ P\left(T_v > t_{\alpha,v} - \frac{\delta}{S_p \sqrt{\frac{2}{n}}}\right) &= 1 - \beta \\ 1 - P\left(T_v < t_{\alpha,v} - \frac{\delta}{S_p \sqrt{\frac{2}{n}}}\right) &= 1 - \beta \\ -t_{\alpha,v} + \frac{\delta}{S_p \sqrt{\frac{2}{n}}} &= t_{\beta,v} \end{aligned}$$

$$n = 2 \left(\frac{\sigma}{\delta} \right)^2 \left\{ t_{\alpha, \nu} + t_{\beta, \nu} \right\}^2$$

Bu teorem, Cochran ve Cox(1957) tarafından ileri sürülmüştür. Örneklem genişliğini hesaplarken, σ 'nın σ_u üst sınırını veya σ 'nın bağımsız tahminini kullanmak gerekir.

Bu problem için alternatif yaklaşım, aşağıda tanımlanan İki Aşamalı yöntemdir:

1. İlk olarak her iki anakütleden n_1 genişlikli örneklem alınır.
2. $S_{p_1}^2$, bu iki örneklemin birleştirilmiş varyansıdır.
3. $c = \left(\frac{\delta^2}{2} \right) \left\{ t_{\alpha, \nu} + t_{\beta, \nu} \right\}^{-2}$ hesaplanır. ($\nu = 2(n_1 - 1)$)
4. $n = \max \left\{ n_1, \left\lceil \frac{S_{p_1}^2}{c} \right\rceil + 1 \right\}$
5. Gerekliyse her iki anakütleden $(n - n_1)$ kadar daha gözlem alınsın.

$\bar{X}$ ve $\bar{Y}$, n gözlemin tümüne dayanan iki örneklem ortalaması olsun. $\mu_1 - \mu_2 = \delta (> 0)$ 'da $(1 - \beta)$ güç gösteren $H_0 : \mu_1 = \mu_2$ tek yanlı α düzeyli testin kritik alanı;

$$\bar{X} - \bar{Y} > \left\{ t_{\alpha} (2n_1 - 2) \right\} S_{p_1} \sqrt{\frac{2}{n}}, \text{dir.}$$

İki yanlı test için, α yerine $\alpha/2$ koyularak benzer bir çözüm elde edilebilir.

$\sigma_1^2 \neq \sigma_2^2$ **Olduğunda ve Varyanslar Bilinmediğinde μ_1 ve μ_2 ile İlgili Hipotez Testleri**

Stein(1945)'in iki aşamalı yöntemi, Chapman(1950) tarafından iki örneklem durumu için geliştirilmiştir. Bu yöntem, aynı serbestlik derecesiyle iki bağımsız Student t değişkeninin farkı olan test istatistiğini kullanır. $T_1(\nu)$ ve $T_2(\nu)$, her biri ν serbestlik dereceli iki bağımsız t değişkeni olsun.

$$M(\nu) = T_1(\nu) - T_2(\nu)$$

Bu deęiřkenin daęılımının zet tablosu Chapman(1950) tarafından verilir. $m_p(v)$, $M(v)$ 'nin deęeri olsun.

α nemlilik dzeyinde, $H_a : \mu_1 > \mu_2$ alternatifine karřı $H_0 : \mu_1 = \mu_2$ sıfır hipotezinin testi dřnlsn. Ayrıca $\mu_1 - \mu_2 = \delta (> 0)$ olduęunda $(1 - \beta)$ g elde etmek istensin. Buna gre ařaęıdaki yntem gereken testi bulur:

1. İki anaktleden her biri n_0 geniřlięinde iki rneklem alınsın. S_{11}^2 ve S_{21}^2 , bu rneklemelerin varyansları olsun.
2. $c = \delta^2 \left\{ m_\alpha (n_0 - 1) + m_\beta (n_0 - 1) \right\}^{-2}$ hesaplanır.
3. $n_i = \max \left(n_0 + 1, \left[S_{i1}^2 / c \right] + 1 \right), i = 1, 2$
4. Anaktleden ($i=1,2$) $(n_i - n_0)$ kadar daha gzlem alınır.
5.
$$b_i = \frac{1}{n_i} \left\{ 1 + \sqrt{\frac{n_0 (n_i c - S_{i1}^2)}{(n_i - n_0) S_{i1}^2}} \right\}$$

$$a_i = \frac{1 - (n_i - n_0) b_i}{n_0}, i = 1, 2$$
6. X_i , X rastlantı deęiřkenine baęlı rneklem ve Y_i , Y rastlantı deęiřkenine baęlı rneklem olsun. ($i = 1, 2, \dots, n_1$ ve $j = 1, 2, \dots, n_2$)

$$T_1 = \frac{\sum_{i=1}^{n_0} a_1 X_i + \sum_{i=n_0+1}^{n_1} b_1 X_i}{\sqrt{c}}$$

$$T_2 = \frac{\sum_{j=1}^{n_0} a_2 Y_j + \sum_{j=n_0+1}^{n_2} b_2 Y_j}{\sqrt{c}}$$

Gereken test iin kritik alan; $T_1 - T_2 > m_\alpha (n_0 - 1)$ 'dir. İki yanlı test benzer řekilde elde edilebilir. (Desu&Raghavarao, 1990)

σ_1^2 ve σ_2^2 ile İlgili Hipotez Testi

Varyans oranı tahmininin devamı olarak, α önemlilik düzeyinde $H_a : \sigma_1^2 > \sigma_2^2$ tek yanlı alternatifine karşı $H_0 : \sigma_1^2 = \sigma_2^2$ sıfır hipotezi test edilmek istensin. Testin kritik alanı; $\frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2} > F_{(1-\alpha), \nu_1, \nu_2}$ 'dir. $F_{(1-\alpha), \nu_1, \nu_2}$; F_{ν_1, ν_2} değişkeninin tablo değeridir. $\frac{\sigma_1^2}{\sigma_2^2} = \lambda^2 (> 1)$ olduğunda $(1 - \beta)$ güç istenmesiyle, $P\left(\frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2} > F_{(1-\alpha), \nu_1, \nu_2} \mid \frac{\sigma_1^2}{\sigma_2^2} = \lambda^2\right) = 1 - \beta$ koşulu elde edilir.

Örneklem genişliği, bu koşulun sağlanmasıyla tanımlanır. Bu koşul;

$$P\left(F_{\nu_1, \nu_2} > \frac{1}{\lambda^2} F_{(1-\alpha), \nu_1, \nu_2}\right) = 1 - \beta$$

şekline dönüştür. Böylece,

$$F_{(1-\alpha), \nu_1, \nu_2} = \lambda^2 F_{\beta, \nu_1, \nu_2} \quad (2.13)$$

olur. Bu eşitlik n_1 ve n_2 için çözülürse aşağıdaki sonuç elde edilir:

Teorem 2.14: $H_0 : \sigma_1^2 = \sigma_2^2$ sıfır hipotezinin α önemlilik düzeyindeki tek yanlı testi

için $\frac{\sigma_1^2}{\sigma_2^2} = \lambda^2 (> 1)$ olduğunda $(1 - \beta)$ güç istenmesiyle gereken örneklem genişliği

(2.13) eşitliğini sağlayan çözümdür.

(2.13) eşitliğinin çözümü tek değildir. $\nu_1 = k\nu_2$ gibi veya daha fazla belirleyerek $\nu_1 = \nu_2$ gibi bir koşulun kabul edilmesiyle tek çözüm elde edilir.

Standart Normal dağılım değerleriyle F dağılım değerleri tahmin edilerek ve $\nu_1 = k\nu_2$ koyulmasıyla ν_1 ve ν_2 'nin çözümü aşağıdaki gibi elde edilir:

$$\nu_1^* = k\nu_2^* \quad \nu_2^* = \left(\frac{k+1}{2k}\right) \left(\frac{Z_\alpha + Z_\beta}{\ln \lambda}\right)^2 \quad (2.14)$$

Teorem 2.15: α önemlilik düzeyinde, $H_0 : \sigma_1^2 = \sigma_2^2$ sıfır hipotezini tek yanlı test

ederken $\frac{\sigma_1^2}{\sigma_2^2} = \lambda^2 (> 1)$ olduğunda $(1 - \beta)$ güç elde etmek için gereken örneklem

genişliği; $\nu_1 = k\nu_2$ olmak üzere, $\mu_i S$ 'nin bilinip bilinmemesine bağlı olarak $[\nu_i^*] + 1$ veya $[\nu_i^*] + 2$ 'dir.

ν_1 ve ν_2 'nin çözümü (2.14) eşitliğinde verilir. ($i=1,2$)

2.2.2. Bernoulli Dağılımlı İki Anakütleden Örnek Seçimi

X 'in; θ_1 başarı olasılığı ile Bernoulli dağılımına, Y 'nin de θ_2 başarı olasılığı ile Bernoulli dağılımına sahip olduğu varsayalım. X rastlantı değişkenlerinden oluşan $X_1, X_2, \dots, X_{n_1}$ rastsal örnekleme ve Y rastlantı değişkenlerinden oluşan $Y_1, Y_2, \dots, Y_{n_2}$ rastsal örnekleme bağımsız olsun. α önemlilik düzeyinde, $H_a: \theta_1 > \theta_2$ tek yanlı alternatifine karşı $H_0: \theta_1 = \theta_2$ sıfır hipotezi test edilirken $\theta_1 - \theta_2 = \delta (> 0)$ olduğunda $(1 - \beta)$ güce sahip olan bir test kullanmak gerekir.

$\hat{\theta}_1 = \sum_{i=1}^{n_1} \frac{X_i}{n_1}$ ve $\hat{\theta}_2 = \sum_{j=1}^{n_2} \frac{Y_j}{n_2}$ olsun. α düzeyli tahmin;

$$\left\{ \arcsin \sqrt{\hat{\theta}_1} - \arcsin \sqrt{\hat{\theta}_2} \right\} > Z_{\alpha/2} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

kritik alanına sahiptir.

Güç koşulu aşağıdaki gibidir:

$$P \left(\arcsin \sqrt{\hat{\theta}_1} - \arcsin \sqrt{\hat{\theta}_2} > Z_{\alpha/2} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \mid \theta_1 - \theta_2 = \delta \right) = 1 - \beta$$

$$\frac{1}{n_1} + \frac{1}{n_2} = \left\{ \frac{2\delta^*}{Z_\alpha + Z_\beta} \right\}^2 \text{ eşitliğini sağlamak için } n_1 \text{ ve } n_2 \text{ örneklem genişliklerini}$$

belirlemek gerekir. Burada δ^* , $\delta^* = \arcsin \sqrt{\theta_1} - \arcsin \sqrt{\theta_2}$ 'e eşittir.

$$\theta_1 \text{ ve } \theta_2 \text{ ayrı olarak verildiğinde } \frac{1}{n_1} + \frac{1}{n_2} = \left\{ \frac{2\delta^*}{Z_\alpha + Z_\beta} \right\}^2 \text{ eşitliği, } n_1 = kn_2 \text{ gibi bir}$$

koşul kullanılarak çözülebilir.

θ_1 ve θ_2 ayrı ayrı belirtilmediğinde en az derecede uyan şekli alınır ve bu eşitlikte δ^* yerine $\delta^* = \arcsin \sqrt{0.50 + \frac{\delta}{2}} - \arcsin \sqrt{0.50 - \frac{\delta}{2}}$ koyulur.

$$\frac{1}{n_1} + \frac{1}{n_2} = \left\{ \frac{2\delta^*}{Z_\alpha + Z_\beta} \right\}^2 \text{ eşitliğinde tek çözüm elde etmek için } n_1 = kn_2 \text{ koyulur.}$$

Buradan

$$n_1^* = kn_2^*, \quad n_2^* = \frac{(k+1)}{k} \left\{ \frac{Z_\alpha + Z_\beta}{2\delta_{\min}^*} \right\}^2 \quad (2.15)$$

bulunur.

Teorem 2.16: α önemlilik düzeyinde, $\theta_1 - \theta_2 = \delta (> 0)$ olduğunda, $H_0 : \theta_1 = \theta_2$ sıfır hipotezinin tek yanlı testinde $(1 - \beta)$ güç elde etmek için gereken n_1 ve n_2 örneklem genişliklerinin tahmini; $[n_1^*] + 1$ ve $[n_2^*] + 2$ 'dir. n_1^* ve n_2^* , (2.15) eşitliğinde verilmiştir.

Bu durumda Fisher-Exact testi de kullanılabilirdi. Bu yöntem kullanılarak örneklem genişliği tahmini Gail-Gart(1973) ve Haseman(1978) tarafından geliştirilmiştir.

2.3.K Örneklem Durumunda Örneklem Genişliği Belirleme

$X_1, X_2, \dots, X_k$ bağımsız rastlantı değişkenlerinin $\mu_i = E(X_i)$ ortalama ve $\sigma_i^2 = V(X_i)$ varyans ile $F_i(x)$ dağılım fonksiyonuna sahip olduğu varsayalım.

Bu anakütleden rastgele olarak çekilen örneklemelere dayanarak $\mu_i - \mu_j$ ($i = j = 1, 2, \dots, k$) ortalamalar farkı tahmini yapılmak ya da μ_i 's ($i = 2, 3, \dots, k$) ile μ_1 karşılaştırılmak istenebilir. Bu karşılaştırma; X_1 kontrol tedavisi için bağımlı değişken ve $X_2, X_3, \dots, X_k$ çalışmadaki $(k-1)$ etkin tedavi için bağımlı değişkenler olduğunda yapılır. İlgilenilen ortalama ile diğer ortalamalar arasındaki fark olan $(\mu_i - \mu_j)$ ($i = j = 1, 2, \dots, k$) incelenir. Bu işlem $(k-1)$ tane tahmin yapmayı gerektirir.

Ayrıca, ilgilenilen olayın belirli alternatifindeki gücün kontrol edilmesiyle $H_0 : F_1 = F_2 = \dots = F_k$ hipotezi test edilmek istenebilir. Bu test problemi, verilen bazı parametre ailelerine uyan F_i 's varsayımıyla incelenecektir.

2.3.1.Ortalamlar Farkının Tahmini

Ortalamlar Farkının Tahmini; anakütle dağılımına ilişkin varsayımlar olmama ve normal dağıldığı varsayımı altında olmak üzere iki alt bölümde incelenecektir.

Dağılıma Ait Varsayımlar Olmadan Tahmin

X_i rastlantı değişkenlerinde oluşan $(i = 1, 2, \dots, k)$ n_i genişlikli rastsal örneklem $X_{i1}, X_{i2}, \dots, X_{in_i}$ ve bu örneklemin ortalaması $\bar{X}_i = \sum_{j=1}^{n_i} \frac{X_{ij}}{n_i}$ olsun. Bu durumda $\mu_i - \mu_j$ farkının yansız tahmin edicisi olarak genellikle $\bar{X}_i - \bar{X}_j$ kullanılır. $k(k-1)/2$ tane fark olduğundan tahmin ediciler için en küçük ortalama varyanslı bütün farklar tahmin edilmek istenebilir. A; tüm $k(k-1)/2$ tahmin edicinin ortalama varyansı olmak üzere aşağıda verilmiştir:

$$A = \frac{2}{k(k-1)} \sum_{i,j=1}^k V(\bar{X}_i - \bar{X}_j) = \frac{2}{k} \sum_{i=1}^k \frac{\sigma_i^2}{n_i}$$

Bu, $c_0 + \sum_{i=1}^n c_i n_i = C$ doğrusal maliyet fonksiyonu kısıtı kullanılarak minimum yapılır.

Cauchy-Schwartz eşitsizliği kullanılarak aşağıdaki sonuç bulunur:

Teorem 2.17: $c_0 + \sum_{i=1}^n c_i n_i = C$ doğrusal maliyet fonksiyonuna bağlı, $\mu_i - \mu_j$ parametrelerinin $k(k-1)/2$ farkının tahmin edicisinin en küçüklenen ortalama varyansı A'nın örneklem genişliği n_i ; $[n_i^*]$ 'ye eşittir. n_i^* aşağıda verilmiştir:

$$n_i^* = \frac{C - c_0}{\sum_{i=1}^k \sigma_i \sqrt{c_i}} \cdot \frac{\sigma_i}{\sqrt{c_i}}, i = 1, 2, \dots, k$$

$(k-1)$ tane $\mu_i - \mu_1 (i = 2, 3, \dots, k)$ farkı tahmin edilmek istendiğinde ortalama varyans A_0 ; aşağıdaki formülle bulunur:

$$A_0 = \frac{1}{k-1} \sum_{i=2}^k V(\bar{X}_i - \bar{X}_1) = \frac{1}{k-1} \left[\frac{(k-1)\sigma_1^2}{n_1} + \sum_{i=2}^k \frac{\sigma_i^2}{n_i} \right]$$

$c_0 + \sum_{i=1}^n c_i n_i = C$ doğrusal maliyet fonksiyonu kısıtına bağlı olarak A_0 en

küçüklenerek aşağıdaki teorem elde edilir:

Teorem 2.18: Parametreler arasındaki $(k-1)$ farkın tahmin edicisinin en küçük ortalama varyansı A_0 'ın örneklem genişliği n_i ; verilen maliyete bağlı olarak $[n_i^*] + 1$ 'e eşittir.

$$n_1^* = \frac{(C - c_0)}{\left\{ \sigma_1 \sqrt{(k-1)c_1} + \sum_{i=2}^k \sigma_i \sqrt{c_i} \right\}} \sqrt{k-1} \cdot \frac{\sigma_1}{\sqrt{c_1}}$$

$$n_i^* = \frac{C - c_0}{\left\{ \sigma_1 \sqrt{(k-1)c_1} + \sum_{i=2}^k \sigma_i \sqrt{c_i} \right\}} \cdot \frac{\sigma_i}{\sqrt{c_i}}, i = 2, 3, \dots, k$$

Bu teoremden, $\sigma_1^2 = \sigma_2^2 = \dots = \sigma_k^2$ ve $c_1 = c_2 = \dots = c_k$ olduğunda $n_1^* / \sqrt{k-1} = n_2^* = \dots = n_k^*$ olduğu görülür. Diğer bir deyişle, anakütle varyansları eşit olduğunda ve her birimden toplanan verinin maliyeti anakütlenin bütünü için aynı olduğunda, X_1 diğer örneklemlemlerle karşılaştırılarak en büyük örneklem genişliği alınır.

$(k-1)$ tane $\mu_i - \mu_1$ ($i = 2, 3, \dots, k$) farkı tahmin edilirken, $\bar{X}_i - \bar{X}_1$ tahmin edicisine bağlı güven elipsoidinin aralığı daraltılabilir. Bu daraltma, $\bar{X}_i - \bar{X}_1$ tahmin edicisinin en küçüklenen genelleştirilmiş varyansıyla yapılır. Genelleştirilmiş varyans aşağıdaki gibi bulunur:

$$V = \det \left(\frac{\sigma_1^2}{n_1} J_{k-1, k-1} + D \left(\frac{\sigma_2^2}{n_2}, \frac{\sigma_3^2}{n_3}, \dots, \frac{\sigma_k^2}{n_k} \right) \right)$$

Bu formül basitleştirildiğinde aşağıdaki gibi olur:

$$V = \frac{\sum_{i=1}^k \frac{n_i}{\sigma_i^2}}{\prod_{i=1}^k \frac{n_i}{\sigma_i^2}}$$

$\sigma_1^2 = \sigma_2^2 = \dots = \sigma_k^2$ ve $c_1 = c_2 = \dots = c_k$ varsayımıyla, maliyet fonksiyonu kısıtına bağlı olarak aşağıda bulunan n örneklem genişliği ile genelleştirilmiş varyans V , en küçüklenir.

Teorem 2.19: $\sigma_1^2 = \sigma_2^2 = \dots = \sigma_k^2$ ve $c_1 = c_2 = \dots = c_k$ olduğunda $\bar{X}_i - \bar{X}_1$ ($i = 2, 3, \dots, k$) farkının genelleştirilmiş varyansı, $n_1 = n_2 = \dots = n_k$ ise, verilen maliyet fonksiyonuna bağlı olarak en küçüklenir. Ortak değer n ; $[n^*] + 1$ 'e eşittir. n^* , aşağıdaki eşitlikle hesaplanır:

$$n^* = \frac{C - c_0}{k \cdot c_1}$$

Normallik Varsayımı Altında Tahmin

X_i ; μ_i ortalama ve σ^2 varyans ile normal dağılıma sahip olsun ($i = 1, 2, \dots, k$).

Mutlak hatanın kontrolüyle eşanlı olarak $k(k-1)/2$ tane $\mu_i - \mu_j$ farkı tahmini d ve α pozitif sabitler olmak üzere aşağıdaki koşulla ifade edilebilir:

$$P\left(\left|(\bar{X}_i - \bar{X}_j) - (\mu_i - \mu_j)\right| \leq d, 1 \leq i < j \leq k\right) \geq 1 - \alpha \quad (2.16)$$

$n_1 = n_2 = \dots = n_k = n$ eşit örneklem genişliği ve σ^2 bilindiği varsayımıyla bu koşulu sağlayan n belirlenir. Yukarıdaki koşul;

$$P\left(\left|Z_i - Z_j\right| \leq d\sqrt{\frac{n}{\sigma^2}}, 1 \leq i < j \leq k\right) \geq 1 - \alpha \quad (2.17)$$

olarak tekrar yazılabilir.

$$Z_i = \frac{(\bar{X}_i - \mu_i)}{\sigma/\sqrt{n}} \quad (i = 1, 2, \dots, k); \text{ bağımsız } N(0, 1) \text{ değişkenleridir.}$$

(2.17) koşulunun sağlanması için;

$$P\left(Z_{\max} - Z_{\min} \leq d\sqrt{\frac{n}{\sigma^2}}\right) \geq 1 - \alpha$$

olması gerekir. Buradan;

$$d\sqrt{\frac{n}{\sigma^2}} = q_{k,\infty}^\alpha$$

eşitliği yazılır. $q_{k,\infty}^\alpha$; Harter(1960) tarafından tablolanan, k bağımsız $N(0, 1)$ değişkenlerinin aralığının üst α yüzdeliğidir.

$$d\sqrt{\frac{n}{\sigma^2}} = q_{k,\infty}^\alpha \text{ eşitliği çözüldüğünde; } n^* = \left\{ \frac{\sigma q_{k,\infty}^\alpha}{d} \right\}^2 \text{ elde edilir.}$$

Teorem 2.20: $P\left(\left|(\bar{X}_i - \bar{X}_j) - (\mu_i - \mu_j)\right| \leq d, 1 \leq i < j \leq k\right) \geq 1 - \alpha$ koşulu ile belirtilen mutlak hatayı kontrol etmek için her bir anakütleden gereken ortak örneklem genişliği; $[n^*] + 1$ 'e eşittir. n^* , yukarıda verilmiştir.

$(k-1)$ tane $\mu_i - \mu_1$ farkı tahmin edilmek istendiğinde mutlak hatayı kontrol etmek için ortak örneklem genişliği;

$$P\left(\left|(\bar{X}_i - \bar{X}_1) - (\mu_i - \mu_1)\right| \leq \delta, 2 \leq i \leq k\right) \geq 1 - \alpha \quad (2.18)$$

koşulunu sağlayacak şekilde seçilmelidir.

Bu koşul sadeleştirilerek aşağıdaki gibi tekrar yazılabilir:

$$P\left(\max_{2 \leq i \leq k} |(Z_i - Z_1)/\sqrt{2}| \leq \delta\sqrt{\frac{n}{2\sigma^2}}\right) \geq 1 - \alpha$$

Bu koşul $\delta\sqrt{\frac{n}{2\sigma^2}}$ yerine Dunnett(1955) tarafından hazırlanmış tablodan $|d|_{k-1,\infty}^\alpha$

değerinin koyulmasıyla çözülür. Bu değer $\max_{2 \leq i \leq k} \{|Z_i - Z_1|/\sqrt{2}\}$ değişkeninin dağılımının

üst 100α yüzdeliğidir.

Teorem 2.21: $P\left(\left|\left(\bar{X}_i - \bar{X}_1\right) - (\mu_i - \mu_1)\right| \leq \delta, 2 \leq i \leq k\right) \geq 1 - \alpha$ ile belirtilen mutlak hatayı kontrol etmek için gereken ortak örneklem genişliği n ; $\lceil n^* \rceil + 1$ 'e eşittir ve

$$n^* = 2 \left\lceil \frac{\sigma |d|_{k-1, \infty}^\alpha}{\delta} \right\rceil^2, \text{ dir.}$$

σ bilinmediğinde, Bölüm 2.2.1'de tartışılan, iki aşamalı metoda benzer Stein(1945) modelini kullanmak gerekir.

2.3.2. Normal Dağılımlı Bir Anakütleden Örnek Seçimi

X_i ; μ_i ortalama ve σ^2 varyansla normal dağılımlı olduğu varsayılın ($i = 1, 2, \dots, k$). $X_{i1}, X_{i2}, \dots, X_{in}$; X_i rastlantı değişkenlerinden oluşan k tane bağımsız rastsal örneklem olsun ($i = 1, 2, \dots, k$).

Bu durumda genellikle σ^2 'nin bilinmediği varsayımı altında aşağıdaki hipotez test edilir:

$$H_0 : \mu_1 = \mu_2 = \dots = \mu_k$$

$$H_a : \text{Bütün } \mu_i \text{ 'ler eşit değildir.}$$

$\bar{X}_i$ ve S_i^2 ; i . örneklem ortalaması ve varyansı, S_p^2 ise birleştirilmiş varyans olsun. Test için kritik alan aşağıdaki gibidir:

$$F^* = \frac{\sum_{i=1}^k n_i (\bar{X}_i - \bar{X}_{..})^2}{(k-1)S_p^2} > F_{1-\alpha, k-1, n-k}$$

Burada $n = \sum_{i=1}^k n_i$ ve $\bar{X}_{..} = \sum_{i=1}^k n_i \bar{X}_i / n$ 'ya eşittir.

Verilen μ_i^* , μ_i ($i = 1, 2, \dots, k$) 'ye eşit olduğunda, testin en az $(1 - \beta)$ güce sahip olması için örneklem genişliği tanımlanır. F^* ; $\mu_i = \mu_i^*$ ($i = 1, 2, \dots, k$) alternatif hipotezi altında, $(k-1)$ ve $(n-k)$ serbestlik dereceli Merkezi Olmayan F değişkeni gibi dağılır. Bu dağılım aşağıdaki eşitlikle bulunan Merkezi Olmama parametresine sahiptir:

$$\Phi = \sqrt{\frac{\sum_{i=1}^k n_i (\mu_i^* - \bar{\mu}^*)^2}{k\sigma^2}}$$

Bu eşitlikte $\bar{\mu}^*$ yerine aşağıdaki eşitlik yazılır:

$$\bar{\mu}^* = \sum_{i=1}^k \frac{n_i \mu_i^*}{n}$$

Bu dağılım, $F_{k-1, n-k, \Phi}$ şeklinde gösterilir. Bu durumda güç koşul aşağıdaki gibi olur:

$$P(F^* > F_{1-\alpha, k-1, n-k}) \geq 1 - \beta$$

Bu koşul aşağıdaki şekilde tekrar yazılabilir:

$$F_{1-\beta, \sum_{i=1}^k n_i - k, \Phi} \geq F_{1-\alpha, k-1, \sum_{i=1}^k n_i - k}$$

$F_{1-\beta, \dots}$; merkezi olmayan dağılımın $(1 - \beta)100$ yüzdeliğidir.

Bu eşitlik, $n_i S$ tanımlamak için çözülmelidir. Bu problemde tek çözüm elde etmek için maliyet ve diğer faktörlere dayandırılan önceden seçilmiş u_i sabitlerine bağlı olarak $n_i = u_i m$ varsayılır. m tanımlandıktan sonra gereken örneklem genişliği bulunur.

$n_1 = n_2 = \dots = n_k = n$ özel durumu için yukarıdaki eşitlik;

$$F_{1-\beta, k-1, k(m^*-1); \Phi} \geq F_{1-\alpha, k-1, k(m^*-1)} \quad (2.19)$$

durumuna indirgenir. Bu eşitlik, güç eğrisi kullanılarak çözülebilir.

Teorem 2.22: $H_0 : \mu_1 = \mu_2 = \dots = \mu_k$ hipotezinin α düzeyli testi için ortak örneklem genişliği m ; $\lceil m^* \rceil + 1$ 'e eşittir. m^* ; (2.19) eşitsizliğini sağlayan pozitif en küçük tam sayıdır.

Örneklem genişliğinin bulunmasında karşılaşılan asıl sorun, μ_i^* 'nin belirtilmesi ve merkezi olmama parametresi Φ 'nin hesaplanmasıdır. Φ , σ^2 sorunlu parametreyi

içerir. μ_i^* , μ ortalamalı normal dağılımın yüzdeliği olarak ifade edilirse, Φ 'nin sorun parametresi σ^2 'den bağımsız olacağı Guenther(1977) tarafından gösterilmiştir. Alternatif olarak, iki ortalama arasında en az Δ fark olması istenebilir. F testinin gücü, Φ fonksiyonunun yükseltilmesiyle meydana geldiğinden iki ortalama arasındaki fark Δ olacak şekilde Φ en küçüklenebilir. Bu, Δ uzaklıkta iki μ_i^*S alınmasıyla ve bu iki ortalamanın ortalamasına eşit olacak μ_i^*S kalmasıyla elde edilebilir.

En küçük Φ sonucu aşağıda verilmiştir:

$$\Phi_{\min} = \sqrt{\frac{m}{2k}} \cdot \frac{\Delta}{\sigma}$$

$\Phi_{\min}$, m 'nin bir fonksiyonudur. Örneklem genişliği iterasyonla güç eğrisinden saptanır. Bowman ve Kastenbaum(1975) tarafından hazırlanan tablolardan; α , $(1 - \beta)$ ve Δ/σ belirlendiğinde örneklem genişliği bulunmuş olur.

İntegralsiz serbestlik dereceli merkezi F dağılımı ile merkezi olmayan F dağılımını tahmin ederek örneklem genişliğini belirlemek için Guenther(1979) tarafından bir yöntem geliştirilmiştir.

F testi yerine; σ^2 bilindiğinde değişim aralığı testi, σ^2 bilinmediğinde Studentlaşmış değişim aralığı testi kullanılabilir.

σ^2 bilinmediğinde alternatiflerin bir grubuna minimum güç sağlamak için iki aşamalı örnekleme planı uygulanabilir. Örneklem ortalamaları ve S_p^2 de (birinci örnekleme dayalı) teste dayandırılır. Bu test ile ilgili detaylar, Hochberg ve Lachenbruch(1976)'ta görülür.

2.3.3. Bernoulli Dağılımlı Bir Anakütleden Örnek Seçimi

X_i 'nin; θ parametrelili Bernoulli dağılımına sahip olduğu varsayılın ($i = 1, 2, \dots, k$). $X_{i1}, X_{i2}, \dots, X_{im_i}, X_i$ rastlantı değişkenlerinden oluşan k tane bağımsız örneklem olsun. Buna göre test edilen sıfır hipotezi;

$$H_0 : \theta_1 = \theta_2 = \dots = \theta_k$$

veya eşiti olarak

$$H_0 : \Phi_1 = \Phi_2 = \dots = \Phi_k$$

olur. Φ_i ; $2 \arcsin \sqrt{\theta_i}$ 'ye eşittir ($i = 1, 2, \dots, k$).

$$H_0 : \Phi_1 = \Phi_2 = \dots = \Phi_k \text{ hipotezi, } H_0 : \sum_{i=1}^k (\Phi_i - \bar{\Phi})^2 \text{ olarak tekrar yazılabilir.}$$

Burada $\bar{\Phi}$, $\Phi_i S$ 'in ortalamasıdır.

Φ_i 'nin tahmin edicisi olarak genellikle $\hat{\Phi}_i = 2 \arcsin(\sqrt{\bar{X}_i})$ kullanılır. $\bar{X}_i$; X_i rastlantı değişkenlerinden oluşan örneklemin ortalamasıdır.

Φ_i 'nin tahmin edicisi olarak genellikle $\hat{\Phi}_i = 2 \arcsin(\sqrt{\bar{X}_i})$ kullanılır. $\bar{X}_i$; X_i rastlantı değişkenlerinden oluşan örneklemin ortalamasıdır. Φ_i 'nin tahmin edicisi $\hat{\Phi}_i$ 'nin yaklaşık olarak $N(\Phi_i, 1/n_i)$ dağıldığı görülür.

$H_0 : \sum_{i=1}^k (\Phi_i - \bar{\Phi})^2$ sıfır hipotezinin α düzeyli testi için kritik alan aşağıda verilmiştir:

$$U = \sum_{i=1}^k n_i (\hat{\Phi}_i - \bar{\Phi})^2 > \chi_{1-\alpha, k-1}^2$$

Burada $\bar{\Phi}$;

$$\bar{\Phi} = \frac{\sum_{i=1}^k n_i \hat{\Phi}_i}{\sum_{i=1}^k n_i}$$

eşittir.

$\Phi_i = \Phi_i^*$ ($i = 1, 2, \dots, k$) olduğunda en az $(1 - \beta)$ güç istenirse $n_i S$ örneklem genişliği;

$$P(U > \chi_{1-\alpha, k-1}^2 \mid \Phi_i = \Phi_i^*) \geq 1 - \beta \quad (2.20)$$

olacak şekilde tanımlanır.

$\Phi_i = \Phi_i^* (i=1,2,\dots,k)$ olduğunda U rastlantı değişkeninin dağılımı, $(k-1)$ serbestlik dereceli merkezi olmayan χ^2 'dir. Merkezi olmama parametresi λ , aşağıdaki gibi bulunur:

$$\lambda = \sqrt{\sum n_i (\Phi_i^* - \bar{\Phi}^*)^2}$$

$$\bar{\Phi}^* = \frac{\sum_{i=1}^k n_i \Phi_i^*}{\sum_{i=1}^k n_i}$$

λ merkezi olmama parametrelili ve $(k-1)$ serbestlik dereceli χ^2 değişkeni, $\chi^2_{k-1;\lambda}$ ile gösterilir. (2.20) koşulu,

$$P(\chi^2_{k-1;\lambda} > \chi^2_{1-\alpha,k-1}) \geq 1 - \beta \quad (2.21)$$

olarak tekrar yazılabilir ve sonuç olarak bu koşulun eşiti aşağıda verilmiştir:

$$\chi^2_{1-\beta,k-1;\lambda} \geq \chi^2_{1-\alpha,k-1} \quad (2.22)$$

$\chi^2_{1-\beta,\dots}$; merkezi olmayan χ^2 dağılımının $100(1-\beta)$ yüzdeliğidir. Bu eşitlik merkezi olmayan χ^2 tablosu kullanılarak λ için çözülür ve örneklem genişliği hesaplanan λ 'dan belirlenir.

Teorem 2.23: $H_0: \theta_1 = \theta_2 = \dots = \theta_k$ sıfır hipotezinin α düzeyli testi için $n_1, n_2, \dots, n_k$ örneklem genişliği, $\chi^2_{1-\beta,k-1;\lambda} \geq \chi^2_{1-\alpha,k-1}$ koşulunu sağlayan pozitif en küçük tam sayıdır.

Genelde Φ_i^* değerinin her birini belirlemek mümkün değildir. Φ_i değerleri arasında en az Δ fark bulmak istenir. $n_1 = n_2 = \dots = n_k = m$ olduğunda Δ 'ya bağlı λ 'nın en küçük değeri aşağıda verilmiştir:

$$\lambda_{\min} = \frac{\Delta\sqrt{m}}{2}$$

(2.22) koşulunu sağlamak için belirlenen m, $\lambda = \lambda_{\min}$ eşitliğini sağlamalıdır.

2.4.Sonlu Anakütleden Örnek Seçimi

Sonlu anakütleden örnek çekilirken araştırmacı, en çok ortalama, toplam ve oran gibi anakütle parametrelerinin tahmini ile ilgilenir. Hipotez testleri, bu tür örneklem çekmede yarar sağlamaz. Parametre tahmini için Bölüm 2.1, 2.2 ve 2.3'te tartışılan yöntemler, sonlu anakütle durumunda uygulanamaz. Çünkü tahmin edicinin dağılımı genellikle bilinmez. Tahmin edicinin değişim aralığı katsayısının veya varyansın kontrolüyle yapılan tahminde, mutlak hatanın veya nisbi hatanın kontrol edildiği görülür. Bu bölümde tahmin edicinin varyansı kontrol edilecektir. Böylece gereken örneklem genişliği belirlenir.

2.4.1.Basit Rastgele Örneklem

N genişlikli anakütleden yerine koymadan n büyüklüğünde basit rastgele örneklem çekildiğini varsayalım. $\bar{x}$; örneklem ortalaması, $\bar{X}$ ve S^2 ; anakütle ortalaması ve varyansı olsun. Bağımlı değişken 0 veya 1 değeri alınarak ikiye ayrıldığında, $\bar{x}$ ve $\bar{X}$; sırasıyla p ve P ile gösterilecektir. Bu durumda, S^2 aşağıdaki gibi olur:

$$S^2 = NPQ / (N-1), Q = 1 - P$$

Bilindiği gibi varyans, $V(\bar{x}) = \left(\frac{N-n}{Nn} \right) S^2$ 'dir.

Önceden seçilmiş V^* 'ye göre, $V(\bar{x}) \leq V^*$ kısıtının kabul edilmesiyle, bu eşitsizliği sağlayacak n belirlenir. Gereken n, aşağıda verilmiştir:

$$n \geq \frac{NS^2}{S^2 + NV^*} = n^* \quad (2.23)$$

Bu eşitlik aşağıdaki teoremle saptanır:

Teorem 2.24: N genişlikli sonlu anakütleden yerine koymadan rastgele örnek çekilirken, $V(\bar{x}) \leq V^*$ koşulu ile istenen örneklem genişliği n; $[n^*] + 1$ 'dir ve

$$n^* = \frac{NS^2}{S^2 + NV^*} \text{ 'dir.}$$

İspat:

$$V(\bar{x}) \leq V^*$$

$$\frac{N-n}{Nn} \cdot S^2 \leq V^*$$

$$\frac{1}{n} - \frac{1}{N} \leq \frac{V^*}{S^2}$$

$$\frac{1}{n} \leq \frac{NV^* + S^2}{S^2 N}$$

$$n^* = \frac{NS^2}{S^2 + NV^*}$$

S^2 ile ilgili önsel bilgi, Teorem 2.24'ü kullanarak örneklem genişliği belirlemek için gereklidir. Bu bilgi, (a) önceki bilgilerden veya (b) pilot anketlerden elde edilebilir. S^2 ile ilgili önsel bilginin kabul edilebilirliği imkansız ise, aşağıdaki iki aşamalı yöntem kullanılabilir:

1. İlk olarak örneklem genişliği $n_1 (\geq 2)$ alınır. bu örneklem varyansı S_1^2 hesaplanır. S_1^2 , S^2 'nin yansız tahmin edicisidir.

$$2. \quad n_2 = \frac{S_1^2}{V^*} \left(1 - \frac{2}{n_1} \right)$$

$$n = \max \left\{ n_1, \left\lceil n_2 \frac{1}{\left(\frac{n_2}{N} \right) + 1} \right\rceil - 1 \right\}$$

3. Gerekirse $n - n_1$ kadar daha gözlem alınır. n gözlemin hepsine dayanan $\bar{x}$ örneklem ortalaması, $\bar{x}$ 'in istenen tahmin edicisidir.

Bağımlı değişken; 0 veya 1 değerleri alınarak ikiye ayrıldığında, n^* için (2.23) eşitliği aşağıda verilen halini alır:

$$n_p^* = \frac{N}{1 + \frac{(N-1)V^*}{PQ}}$$

P genellikle bilinmediğinden ve tahmin edilmiş parametre için tahmin yapmak oldukça zor olduğundan P kendisidir. İlimli bir yaklaşım olarak $P = \frac{1}{2}$ alınmaktadır.

Teorem 2.25: $V(p) \leq V^*$ olacak şekilde P anakütle oranını tahmin etmek için gereken örneklem genişliği n, $[n_{0.5}^*] + 1$ 'e eşittir. n_p^* , yukarıda verilmiştir.

İspat:

$$V(p) \leq V^*$$

$$\frac{N-n}{Nn} \cdot \frac{NPQ}{N-1} \leq V^*$$

$$\frac{1}{n} - \frac{1}{N} \leq \frac{V^* \cdot (N-1)}{NPQ}$$

$$n_p^* = \frac{N}{1 + \frac{(N-1)V^*}{PQ}}$$

Aşağıdaki bölümde, kabul örnekleme bağlamında örneklem genişliği belirleme tartışılacaktır.

Kabul Örnekleme

İncelenen anakütle, kusurlu olan $D = NP$ parçasını oluşturan N'den meydana gelir. Kalan $N - D = NQ$ parça iyidir. Genellikle D bilinmez. n genişlikli örneklem sorunda $H_a : D > D_0$ alternatifine karşı $H_0 : D = D_0$ sıfır hipotezini test etmek istenir. α düzeyli bu testin kritik alanı; $Y > C$ 'dir. Y; örneklemdeki kusurlu parçaların sayısıdır. Pozitif tam sayı C, aşağıdaki gibi belirlenir:

$$\sum_{y=C+1}^{\min(D_0, n)} \frac{\binom{D_0}{y} \binom{N-D_0}{n-y}}{\binom{N}{n}} \leq \alpha$$

$D = D_1 (> D_0)$ olduğunda, $(1 - \beta)$ 'dan küçük olmayan bir güce sahip olmak istenir. Bu durumda koşul, aşağıdaki gibidir:

$$\sum_{y=c+1}^{\min(D_0, n)} \frac{\binom{D_1}{y} \binom{N-D_1}{n-y}}{\binom{N}{n}} \geq 1 - \beta$$

Yukarıdaki iki eşitlik n için çözülebilir. Bu eşitliklerde kullanılan olasılıklar, hipergeometrik dağılıma uyar ve Lieberman ve Owen(1961) tarafından tablo haline getirilmiştir. Bu probleme göre tahmini çözüm, Guenther(1977) tarafından incelenmiştir.

$N(nP, nPQ)$ ile Y 'nin dağılımı tahmin edilebilir ve aşağıdaki sonuç elde edilir:

Teorem 2.26: $D = D_1 (> D_0)$ olduğunda en az $(1 - \beta)$ güç elde etmek için $H_a : D > D_0$ tek yanlı alternatifine karşı $H_0 : D = D_0$ sıfır hipotezinin α önemlilik düzeyindeki testi için gereken örneklem genişliği $n; [n^*] + 1$ ile tahmin edilir.

$$n^* = \left\{ \frac{Z_\alpha \sqrt{D_0(N-D_0)} + Z_\beta \sqrt{D_1(N-D_1)}}{(D_1 - D_0)} \right\}^2$$

2.4.2. Tabakalı Rastgele Örneklem

L katmana tabakalanan anakütlenin h . Tabakasının N_h genişliğinde olduğu ve anakütle varyansının S_h^2 ($h=1, 2, \dots, L$) olduğu düşünölsün. Araştırmacı, n genişlikli tabakalı rastgele örnekleme uygularsa, tabakalar arasında bu toplam n genişliğini parçalamaya karar vermek zorundadır. n_h , h tabakaya ayrılmış örneklem genişliği olsun ($h=1, 2, \dots, L$).

$\bar{X}$; anakütle ortalaması, $\bar{x}_h$; h . Tabakadan örneklem ortalaması olsun.

$$\hat{\bar{X}} = \sum_{h=1}^L \frac{N_h \bar{x}_h}{N}, \quad N = \sum_{h=1}^L N_h \text{ olduğu biliniyor.}$$

$$V(\hat{\bar{X}}) = \frac{1}{N^2} \sum_{h=1}^L N_h (N_h - n_h) \cdot \frac{S_h^2}{n_h} \quad (2.24)$$

$n_h = \frac{nN_h}{N}$ ise, toplam örneklem genişliğinin tabakalara ayrılması, oransal ayırma

olarak adlandırılır. Bu durumda,

$$V_{oran}(\hat{\bar{X}}) = \frac{N-n}{N^2n} \sum_{h=1}^L N_h S_h^2$$

olur.

$V_{oran}(\hat{\bar{X}}) \leq V^*$ koşulunun kabul edilmesiyle bu varyans kontrol edilmek

istendiğinde gereken toplam örneklem genişliği n ; $[n^*] + 1$ 'e eşit olur.

Teorem 2.27: Oransal ayırma ile $V_{oran}(\hat{\bar{X}}) \leq V^*$ koşulunu sağlamak için gerekli olan

toplam örnekleme genişliği, $[n^*] + 1$ 'e eşittir. $n^*, n^* = N \left\{ \frac{\sum_{h=1}^L N_h S_h^2}{V^* N^2 + \sum_{h=1}^L N_h S_h^2} \right\}$ 'dir.

Ayrıca, h . Tabaka için örneklem genişliği; $[n^* n_h / N] + 1$ 'e eşit olur ($h = 1, 2, \dots, L$).

İspat:

$$V_{oran}(\hat{\bar{X}}) \leq V^*$$

$$\frac{N-n}{N^2n} \sum_{h=1}^L N_h S_h^2 \leq V^*$$

$$\frac{1}{Nn} - \frac{1}{N^2} \leq \frac{V^*}{\sum_{h=1}^L N_h S_h^2}$$

$$n^* = N \left\{ \frac{\sum_{h=1}^L N_h S_h^2}{V^* N^2 + \sum_{h=1}^L N_h S_h^2} \right\}$$

Oransal ayırmayı kullanmak yerine, $c_0 + \sum_{h=1}^L c_h n_h = C$ doğrusal maliyet

fonksiyonunu kullanarak maliyet kısıtına bağlı $V(\hat{\bar{X}})$ için en uygun ayırma kullanılmak

istenebilir. Gerekli n_h değeri, (2.24) eşitliğinde N_h ile $N_h - n_h$ tahmininden sonra

Cauchy-Schwartz eşitliğinden elde edilebilir ve aşağıdaki teoremden belirtilen eşitlik bulunur:

Teorem 2.28: $C = c_0 + \sum_{h=1}^L c_h n_h$ maliyet fonksiyonuna bağlı $V(\hat{\bar{X}})$ 'i en küçükleyen n_h değeri, $[n_h^*] + 1$ 'e eşittir ($h = 1, 2, \dots, L$). n_h^* eşitliği aşağıda verilmiştir:

$$n_h^* = \frac{(C - c_0) N_h S_h / \sqrt{c_h}}{\sum_{h=1}^L N_h S_h / \sqrt{c_h}}, \quad (h = 1, 2, \dots, L)$$

Benzer sonuçlar, anakütle oranı tahmin edilirken de elde edilebilir.

2.4.3. Tabakalamaya Göre Tam Örnekleme

Toplam anakütle genişliği N olan bir anakütle düşünelim. Bu anakütle; $N_1, N_2, \dots, N_L$ genişlikli L tabakaya bölünsün. Araştırmacı, bu tabakaların genişlikleri bilgisine sahip değildir. Bu durumda ilk olarak n' genişliğinde basit rastgele örneklem çekilir. Bu örnekleme dayanarak araştırmacı, $N n_h' / n'$ ile tabaka genişliği tahmini yapar. n_h' ; bu örnekleme h tabakaya ayrılan birimlerin sayısıdır. İkinci aşama olarak, birinci aşamadaki n_h' birimden, $n_h = v_h n_h'$ örneklemini rastgele çekilir. $\bar{x}_h$, ikinci aşamadaki h . Tabaka örneğin ortalaması olmak üzere anakütle ortalaması $\bar{X}$; aşağıdaki eşitlikle tahmin edilir:

$$\hat{\bar{X}} = \sum_{h=1}^L \frac{n_h'}{n'} \bar{x}_h$$

$V(\hat{\bar{X}})$ tahmini ise;

$$V(\hat{\bar{X}}) = S^2 \left(\frac{1}{n'} - \frac{1}{N} \right) + \sum_{h=1}^L \frac{N_h S_h^2}{N n'} \left(\frac{1}{V_h} - 1 \right)$$

ile bulunur.

Tahmin edilen varyans; $C = c_0 + c' n' + \sum_{h=1}^L c_h n_h$ maliyet fonksiyonu kullanılarak

Cauchy-Schwartz eşitsizliği ile gösterilen beklenen maliyetin bulunmasıyla en küçüklenir:

$$E(C) = c_0 + c' n' + n' \sum_{h=1}^L c_h v_h \frac{N_h}{N} = C^* \quad (2.25)$$

$$v_h = S_h \left\{ \frac{c'}{c_h \left(S^2 - \sum_{h=1}^L \frac{N_h S_h^2}{N} \right)} \right\}^{\frac{1}{2}} \quad (2.26)$$

Birinci aşamadaki örneklem genişliği n' , (2.25) eşitliğinde v_h yerine koyularak

çözülebilir. n' ; $[n'] + 1$ 'e eşittir ve $n' = \frac{C^* - c_0}{c' + \sum_{h=1}^L c_h V_h N_h / N}$ 'dir.

Teorem 2.29: Tabakalamaya göre tam örneklemede, birinci aşamada basit rastgele örneklemin tahmini genişliği, $[n'] + 1$ 'e eşittir. n' eşitliği yukarıda verilmiştir. n'_h , h. Tabakaya ait birinci aşama birimleri ise h. Tabakadan ikinci aşama seçilen birimlerin sayısı, $[n'_h v_h] + 1$ 'e eşittir. v_h , (2.26) eşitliği ile bulunur.

2.4.4.İki Aşamalı Küme Örnekleme

Anakütle N tane kümeden oluşsun (Birinci aşama örnekleme birimleri olarak da bilinsin). M_i ; i. kümede alt birimlerin sayısı olsun (İkinci aşama birimleri olarak da bilinsin) ($i = 1, 2, \dots, N$). Y_{ij} ; i. küme j. Aşama birimi olsun.

$$Y_i = \sum_{j=1}^{m_i} Y_{ij}, \quad \bar{Y}_i = Y_i / m_i, \quad Y = \sum_{i=1}^N Y_i, \quad \bar{Y} = Y / N$$

$$S_{w_i} = \sum_{j=1}^{m_i} (Y_{ij} - \bar{Y}_i)^2 / (M_i - 1) \text{ olsun.}$$

Z_i için oransal olasılıklarla n küme örnekleme yerine koyularak seçilir. Seçilen i . küme içinde ikinci aşama birim örnekleme m_i , rastgele örnekleme düzeni kullanılarak yerine koymadan seçilir. $\bar{y}_i$ ($i=1,2,\dots,n$); i . küme örnekleme, ikinci aşama birimine bağımlı örnekleme ortalaması olsun. Toplam anakütlede Y 'nin yansız tahmin edicisi aşağıda verilmiştir:

$$\hat{Y} = \sum_{i=1}^n \frac{M_i \bar{Y}_i}{n Z_i}$$

$\hat{Y}$ 'nin varyansı,

$$V(\hat{Y}) = \frac{1}{n} \left[V_z + \sum_{i=1}^N \frac{1}{Z_i} M_i^2 \left(\frac{1}{m_i} - \frac{1}{M_i} \right) S_{y_i}^2 \right] \quad (2.27)$$

ile bulunur.

$$V_z = \sum_{i=1}^N Z_i \left(\frac{Y_i}{Z_i} - Y \right)^2$$

c_0 ; sabit maliyet, c_1 ; küme örnekleme için maliyet değişkeni, c_2 ; seçilen kümede ikinci aşama birimleri seçmek için maliyet değişkeni, c_3 ; her ikinci aşama örnekleme için maliyet değişkeni olsun. Araştırmanın toplam maliyeti aşağıda verilmiştir:

$$C = c_0 + c_1 n + c_2 \sum_{i=1}^n M_i + c_3 \sum_{i=1}^n m_i$$

Beklenen maliyet;

$$E(C) = c_0 + c_1 n + c_2 n \sum_{i=1}^N Z_i M_i + c_3 n \sum_{i=1}^N z_i m_i \quad (2.28)$$

ile bulunur.

Araştırmanın beklenen sabit maliyeti C^* için,

$$\sqrt{n \left(c_1 + c_2 \sum_{i=1}^N Z_i M_i \right)} \propto \sqrt{\frac{1}{n} \left(V_z - \sum_{i=1}^N \frac{M_i S_{w_i}^2}{Z_i} \right)} \quad (2.29)$$

$$\sqrt{c_3 n z_i m_i} \propto \sqrt{\frac{1}{n} \cdot \frac{M_i^2}{z_i m_i} \cdot S_{w_i}^2}$$

olduğunda Cauchy–Schwartz eşitliği ile $V(\hat{Y})$ en küçüklenir.

(2.29)’dan;

$$m_i^* = \frac{M_i S_{w_i}}{Z_i \sqrt{V_z - \sum_{i=1}^N \frac{M_i S_{w_i}^2}{Z_i}}} \sqrt{\frac{c_1 + c_2 \sum_{i=1}^N Z_i M_i}{c_3}} \quad (2.30)$$

bulunur.

(2.28) eşitliğinin sağ tarafında m_i^* değeri yerine koyulup c^* ’a eşitlenerek n^* elde edilir.

Teorem 2.30: Yerine koyularak ve Z_i için orantısal olasılıklarla seçilen iki aşamalı küme örneklemede, $\hat{Y}$ tahmin edicisinin (2.27)’de bulunan varyansı, seçilen i. kümeden $[m_i^*] + 1$ ikinci aşama birim alınmasıyla en küçüklenir. m_i^* , (2.30) eşitliği ile bulunur. Seçilecek kümelerin sayısı, $[n^*] + 1$ ’e eşittir. n^* eşitliği aşağıda verilmiştir:

$$n^* = \frac{c^* - c_0}{c_1 + c_2 \sum_{i=1}^N Z_i M_i + c_3 \sum_{i=1}^N Z_i m_i^*}$$

Tam örnekleme ve İki aşamalı örnekleme, benzer görünmelerine karşın birbirinden farklıdır. Tam örneklemede, ikinci aşama örneklem birinci aşama örneklemin alt kümesidir. İki aşamalı örneklemede ise, ikinci aşama örneklem kalan anakütleden çekilir.

Tasarı parametrelerinin ilişki değerleri, varyans oranı ve maliyet oranı kesin olarak bilinmediğinde ama belirli aralıklarda olacağı tahmin edildiği zaman bu özel durum, Sadooghi–Alvandi(1986) tarafından incelenmiştir.

3. KLİNİK ÇALIŞMALARDA ÖRNEKLEM SEÇİMİ

Klinik denemelerde örneklem genişliği belirleme, genellikle hipotez testleri ile ilgilidir. Bu deneylerdeki veri toplama, özel problemlerle ilişkilidir. Hastalar çalışmadan ayrılabilirler veya hayvanlar bilinmeyen nedenlerden kaynaklı olarak ölebilirler. Bu faktörler, bazı deneysel birimlerde yapılan gözlemlerde eksikliklere neden olur. Bölüm 2’de gözlemlerin eksiksiz olduğu varsayımıyla örneklem genişliğini belirleyen yöntemlerden bahsedildi. Bu bölümde, 2. bölümde incelenmeyen bazı özel deneysel düzenler tartışılacaktır.

3.1.Nitel Bağımlı Değişkenlerle Deneyler

Bu bölümdeki deneyler; iki tedavinin etkilerini karşılaştırmak amaçlıdır. Bağımlı değişken; niteldir. Olası değerler sınırlıdır ve sadece iki değer olduğunda 0 ve 1 olarak kodlanır. Sıfır hipotezi; iki değişkenin etkilerinin eşit olmasıdır. Öngörölmüş α düzeyli testler bulunur ve örneklem genişlikleri, bu testlerin elde edilen alternatiflerde öngörölmüş güçte olacak şekilde belirlenir.

3.1.1.McNemar Testi

Aşağıda deneysel durumların iki türü incelenir:

- 1) Deneysel birimlerin (insan, hayvan, vs.) eşleştirilmiş çifti n , deneyde kapsanır(içerilir). Her çifte göre birimler, rastgele olarak tedavi A ve tedavi B’ye atanır. Bağımlı her birime göre 0 – 1 atanır.
- 2) n deneysel birimlerin her biri A ve B tedavilerinin her ikisinden elde edilir. Rastgele olarak seçilen n_1 birimi, birbirini izleyen tedavilerden A’yı izleyen B ile bulunur ve kalan $n - n_1$ birimi, B tedavisini izleyen A ile bulunur. Aynı deneyler, çapraz veya etki değişimi tasarımı olarak bilinir. 0 – 1 A bağımlısı, her tedavi altındaki her denek için kodlanır. Tedaviler arasındaki yeterli durgunluk süresi, aktarma etkilerinden kurtulmak için kullanılmalıdır.

Bu durumların her ikisindeki sonuçlar, n_{ij} teriminde toplanır. n_{ij} ; A’ya göre i , B’ye göre j bağımlısı ile çiftlerin (birimlerin) sayısıdır. Özet tablo aşağıdadır:

		Tedavi B	
		0	1
Tedavi A	0	n_{00}	n_{01}
	1	n_{10}	n_{11}

3.1. Tedavi A ve B'ye Göre Birimlerin Sayısı

Göze frekansı n_{ij} ; θ_{ij} ($i, j = 0, 1$) ve n parametrelili çok terimli dağılıma uyar.

İddia edilen hipotez; iki tedavinin etkilerinin eşitliği ifadesiyle yani,

$$H_0 : \theta_{10} + \theta_{11} = \theta_{01} + \theta_{00}$$

veya

$$H_0 : \theta_{10} = \theta_{01}$$

şeklinde kurulur. Alternatif hipotez; $H_a : \theta_{10} > \theta_{01}$ 'dir. Aynı deneyde bir tedavi standart veya kontrol, diğeri ise standarttan daha iyi olması beklenen deneysel tedavidir.

$\frac{n_{10}}{n} - \frac{n_{01}}{n} > c$ kritik alanlı test incelenir. Buradaki c sabiti, testin α düzeyli olması için seçilir. Buradan, c ;

$$P(n_{10} - n_{01} > cn \mid H_0) \leq \alpha$$

koşulunu sağlamalıdır.

$$n_{10} - n_{01} = 2(n_{10} - S/2)$$

olması için

$$S = n_{10} + n_{01}$$

olsun. Ayrıca, n_{10} 'ın dağılımını bulan S ;

$$s \text{ ve } \theta_{10} / (\theta_{10} + \theta_{01}) = \theta_{10} / \psi$$

parametrelili binom dağılımıdır.

$$P(n_{10} - n_{01} > cn \mid H_0) \leq \alpha$$

koşulu her bir S değerine göre,

$$P\left(n_{10} - \frac{S}{2} > \frac{cn}{2} \mid S, H_0\right) \leq \alpha$$

ise sağlanacaktır. n büyük olduğunda, dönüşüm

$$cn/2 = Z_\alpha \sqrt{(S/4)}$$

seçilmesiyle yaklaşık olarak sağlanabilir. α düzeyli test tahmini aşağıda verilen kritik alan ile bulunur:

$$n_{10} > \frac{S}{2} + Z_\alpha \sqrt{\frac{S}{4}}$$

$\theta_{10} - \theta_{01} = \delta^* > 0$ 'a göre $(1 - \beta)$ güç kontrol etmek istenirse,

$$P\left(n_{10} > \frac{S}{2} + Z_\alpha \sqrt{\frac{S}{4}} \mid \theta_{10} - \theta_{01} = \delta^*\right) = 1 - \beta \quad (3.1)$$

olması için n seçilmelidir. Bu eşitliğin sol tarafındaki olasılık aşağıdaki gibi tahmin edilir:

$$E_s P\left(n_{10} > \frac{S}{2} + Z_\alpha \sqrt{\frac{S}{4}} \mid S, \theta_{10} - \theta_{01} = \delta^*\right) = E_s P\left[1 - \Phi\left\{\left(\frac{S}{2} + Z_\alpha \sqrt{\frac{S}{4}} - \frac{S\theta_{10}}{\psi}\right) / \sqrt{S \frac{\theta_{10}}{\psi} \left(1 - \frac{\theta_{10}}{\psi}\right)}\right\}\right]$$

n_{10} koşullu binom dağılımı tahmini ile bulunan S, normal dağılımlıdır.

Dönüşümdeki bu ifade;

$$1 - \Phi\left[\left\{E_1(s) \left(\frac{1}{2} - \frac{\theta_{10}}{\psi}\right) + \frac{Z_\alpha}{2} \sqrt{E_1(s)}\right\} / \sqrt{\frac{\theta_{10}}{\psi} \left(1 - \frac{\theta_{10}}{\psi}\right) E_1(s)}\right] \quad (3.2)$$

ile tahmin edilebilir. $E_1(s)$ eşitliği aşağıda verilmiştir:

$$E_1(s) = E(S \setminus \theta_{10} - \theta_{01} = \delta^*) = n(\theta_{10} + \theta_{01}) = n\psi$$

$(1 - \beta)$ ile (3.2) eşitliğinde bulunan güç tahmininin eşit varsayılmasıyla, (3.1)

eşitliğinin sağlayan n 'in tahminini veren eşitlik aşağıdaki gibidir:

$$n\psi \left(\frac{1}{2} - \frac{\theta_{10}}{\psi} \right) + \frac{Z_\alpha}{2} \sqrt{n\psi} = -Z_\beta \sqrt{\theta_{10} \left(1 - \frac{\theta_{10}}{\psi} \right) n}$$

$$\sqrt{n}\delta^* = Z_\alpha \sqrt{\psi} + Z_\beta \sqrt{\frac{4\theta_{10}\theta_{01}}{\psi}} \text{ eşitliğine benzer olarak,}$$

$$\begin{aligned} \sqrt{n}(2\theta_{10} - \psi) &= Z_\alpha \sqrt{\psi} + Z_\beta \sqrt{4\theta_{10} \left(1 - \frac{\theta_{01}}{\psi} \right)} \\ &= Z_\alpha \sqrt{\psi} + Z_\beta \sqrt{\frac{(\psi + \delta^*)(\psi - \delta^*)}{\psi}} \end{aligned} \quad (3.3)$$

için aynıdır.

n için, (3.3) eşitliğinin çözümü sırasında ψ belirlemek gerektiğine dikkat edilmelidir.

Teorem 3.1.: $\theta_{10} - \theta_{01} = \delta^* > 0$ ve $\theta_{10} + \theta_{01} = \psi$ 'da $(1 - \beta)$ güçle $H_0 : \theta_{10} = \theta_{01}$ tek yanlı testin α düzeyli tahmini için gereken örneklem genişliği tahmini, $[n^*] + 1$ 'dir. n^* , (3.3) eşitliğini sağlar.

Teorem 3.1'de incelenen tahmin, Lachin(1981) tarafından geliştirilmiştir. Bu, çok az farklı bir yaklaşım kullanılarak Lee v.d. (1984) tarafından da elde edilmiştir. Koşulsuz güç fonksiyonuna dayandırılan örneklem genişliği tablosu Schork ve Williams(1980) tarafından bulunur. Bu problem, örneklem genişliğine alternatif bir yaklaşım sunan Connor(1987) tarafından da incelenmiştir.

Alternatif hipotez $H_a : \theta_{10} \neq \theta_{01}$ şeklinde iki yanlı olduğunda McNemar testi için kritik alan aşağıdaki gibi olur:

$$T = \frac{(n_{10} - n_{01})^2}{(n_{10} + n_{01})} > \chi_{1-\alpha,1}^2$$

$|\theta_{10} - \theta_{01}| = \delta^*$ alternatifi için $(1 - \beta)$ 'da güç kontrolü sırasında, n;

$$P(T > \chi^2_{1-\alpha,1} \setminus \delta^*) = 1 - \beta \quad (3.4)$$

olması için seçilmelidir.

$T \sim \chi^2(1, \Delta)$ alternatifi altında; Δ eşitliği aşağıdaki gibidir:

$$\Delta = E(T) - 1 = \frac{n\delta^{*2}}{\psi} - \frac{\delta^{*2}}{\psi}$$

Buradan, (3.4) koşulu,

$$P(\chi^2(1, \Delta) > \chi^2_{1-\alpha,1}) = 1 - \beta \quad (3.5)$$

şeklinde görülebilir.

Merkezi olmayan χ^2 dağılım tablosundan gereken n belirlenir. $N(0,1)$ dağılımıyla,

$$Z = \frac{\sqrt{2\chi^2_{v,\Delta}} - \sqrt{2(v + \Delta^2)} - \frac{v + 2\Delta^2}{v + \Delta^2}}{\sqrt{\frac{v + 2\Delta^2}{v + \Delta^2}}}$$

dağılımının tahmin edilmesiyle, (3.5) koşulunu sağlayan n tahmini, bulunur.

Teorem 3.2: $|\theta_{10} - \theta_{01}| = \delta^*$ ve $\psi = \theta_{10} + \theta_{01}$ 'de $(1 - \beta)$ güçle $H_a : \theta_{10} \neq \theta_{01}$ alternatifine karşı $H_0 : \theta_{10} = \theta_{01}$ sıfır hipotezinin α düzeyli testi için gereken örneklem genişliği

tahmini; $[n^*] + 1$ 'dir. $n^*, n^* = \frac{\psi}{\delta^{*2}} (Z_{\alpha/2} + Z_\beta)^2 + \frac{1}{\psi}$ 'dir.

0-1 bağımlı değişkenli biyoeşdeğerlik problemi Dunnet ve Gent(1977) tarafından incelenmiştir ve gerekli örneklem genişliği tartışması Rodary v.d.(1989) tarafından bulunmuştur.

3.1.2. Kategorik Bağımlılarla (c) Tedavi Etkilerinin (r) Eşitliği Testi

r tedavili bir deney inceleyelim. Burada n_i elemanları i. tedaviden alınıyor ($i = 1, 2, \dots, r$). c bağımlı değişken değerleri alınır. $n = \sum_{i=1}^r n_i$ ve $Q_i = n_i / n$ ($i = 1, 2, \dots, r$) olsun. Veri; rxc tablosunda özetlenir:

		Bağımlı Kategoriler						
		1	2	...	j	...	c	Toplam
Tedaviler	1	n_{11}	n_{12}	...	n_{1j}	...	n_{1c}	n_1
	2	n_{21}	n_{22}	...	n_{2j}	...	n_{2c}	n_2
	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
	i	n_{i1}	n_{i2}	...	n_{ij}	...	n_{ic}	n_i
	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
	r	n_{r1}	n_{r2}	...	n_{rj}	...	n_{rc}	n_r

3.2. Bağımlı Kategoriler ve Tedavilere Göre Birimlerin Sayısı

Her i için $\sum_{j=1}^c \theta_{ij} = 1$ ise θ_{ij} ; i. tedavi altında j. bağımlı elde etme olasılığı olsun.

Sıfır hipotezi, “Bütün tedaviler aynı etkiye sahiptir” şeklindedir ve aşağıdaki gibi gösterilir:

$$H_0 : \theta_{ij} = \alpha_j, \text{ her } i \text{ için } j = 1, 2, \dots, c$$

$$\hat{\alpha}_j = \sum_{i=1}^r n_{ij} / n$$

χ^2 testinin kritik alanı aşağıda verilmiştir:

$$T = \sum_{i,j} \frac{(n_{ij} - n_i \hat{\alpha}_j)^2}{n_i \hat{\alpha}_j} > \chi_{1-\alpha, \nu}^2 \quad \nu = (r-1)(c-1)$$

Belirli δ_{ij} için $\theta_{ij} = \theta'_{ij} = \alpha_j + \delta_{ij}$ parametrelerinin belirli alternatifinde $(1 - \beta)$ güç olması istenirse örneklem genişliği,

$$P\left(T > \chi^2_{1-\alpha, \nu} \setminus \theta'_{ij}\right) = 1 - \beta \quad (3.6)$$

koşulunu sağlayacak şekilde belirlenir.

Alternatif altında T; Δ merkezi olmama parametresi ve ν serbestlik derecesiyle merkezi olmayan χ^2 dağılımına yaklaşıır.

$$\Delta^2 = n \sum_j \frac{1}{\alpha_1} \left\{ \sum_i Q_i \delta_{ij}^2 - \left(\sum_i Q_i \delta_{ij} \right)^2 \right\}, r \geq 2 \quad c \geq 2 \text{ için}$$

$$= n Q_1 Q_2 \left\{ \sum_j \frac{(\delta_{1j} - \delta_{2j})^2}{\alpha_1} \right\}, r = 2 \quad c \geq 2 \text{ için}$$

$$= n \left\{ \sum_i Q_i \delta_{i1}^2 - \frac{\left(\sum_i Q_i \delta_{i1} \right)^2}{\alpha_1 (1 - \alpha_1)} \right\}, r \geq 2 \quad c = 2 \text{ için}$$

Δ^2 merkezi olmama parametresi, payda

$$E(n_{i1} = n_i \theta'_{i1}) \quad E(\hat{\alpha}_j) = \alpha_1 + \sum_i n_i \delta_{ij} / n$$

ve paydada

$$\hat{\alpha}_j = \alpha_j$$

koyulmasıyla bulunur. (3.6) koşulundan,

$$P\left(\chi^2_{\nu, \Delta} > \chi^2_{1-\alpha, \nu}\right) = 1 - \beta$$

bulunur.

Verilen $\{\alpha_j\}$ ve $\{\delta_{ij}\}$ için merkezi olmayan χ^2 dağılım tablosuna bakılarak gereken örneklem genişliği n_i belirlenir.

3.1.3.Uzun Dönem Denemeler

Örneklem genişliği; kontrol tedavisi ve deneysel tedavi, olay oranlarına göre uzun dönem karşılaştırıldığında belirlenir. Çalışmanın uzun süreli olması nedeniyle bazı denekler çeşitli sebeplerden çalışmadan ayrılabilirler. Verilen alternatifte güç kontrol

edilerek olay oranlarının eşitliğinin testi için örneklem genişliği belirlenmek istenir. Örneklem genişliğini belirleme, çalışmadan ayrılan deneklerin olasılığı dikkate alınmadan Bölüm 2’de tartışılmıştı. R; ayrılanların beklenen oranı, n; ayrılmadıkları varsayılarak hesaplanan ortak örneklem genişliği olmak üzere, ayrılan deneklerle gereken örneklem genişliği, $n_d = \frac{n}{(1-R)^2}$ eşitliği ile bulunur.

Halperin v.d.(1968), Wu v.d.(1980)ve Lakatos(1986) bu konuyu daha detaylı incelemiştir.

3.1.4.Olgu–Kontrol Çalışmaları

Olgu–Kontrol çalışmalarında belirli sağlıklı veya hastalıklı (olgular) bireyler ve sağlıklı veya hastalıklı (kontroller) bireyler, çalışma altında, hastalık veya durumlarının gelişiminin risk faktörlerinin etkisi çalışmasına göre karşılaştırılır. Bu bölümde, etki karışımı değişkeninin düzeyleri kullanılarak tabakalanan çalışmanın risk faktörü incelenir.

Etki karışımı değişkeninin derecesine dayanan j. tabakadaki veri ($j = 1, 2, \dots, J$) aşağıdaki tabloda gösterilir:

Risk Grubu	Olgu	Kontrol	Toplam
Etkilenmiş	a_j	b_j	m_{1j}
Etkilenmemiş	c_j	d_j	m_{2j}
Toplam	n_{1j}	n_{2j}	n_j

3.3.Etki Karışımı Değişkeninin Derecesine Dayanan j. Tabakadaki Veriler

$$n = \sum_{j=1}^J n_j t_j = n_j / n \text{ ve } S_j = n_{1j} / n_j \text{ olsun.}$$

Bu durumda j. tabakada maruziyet ile hastalık arasındaki Odds oranı aşağıdaki gibi hesaplanır:

$$\theta_j = \frac{\frac{\pi_{1j}}{(1-\pi_{1j})}}{\frac{\pi_{2j}}{(1-\pi_{2j})}}, (j = 1, 2, \dots, J)$$

$\theta_1 = \theta_2 = \dots = \theta_J (= \theta)$ varsayılır ve $H_a : \theta > 1$ alternatifine karşı $H_0 : \theta = 1$ sıfır hipotezi test edilir. Sıfır hipotezi,

$$\pi_{1j} = \pi_{2j} \quad (j=1,2,\dots,J)$$

şeklindedir.

π_{1j} ; $P_{1j} = a_j / n_{1j}$, π_{2j} ; $P_{2j} = b_j / n_{2j}$ ile tahmin edilir. Cochran(1954) tarafından ileri sürülen test istatistiği aşağıda verilmiştir:

$$C_0 = \frac{\sum_{j=1}^J w_j (P_{1j} - P_{2j})}{\sqrt{\sum_{j=1}^J w_j \bar{P}_j \bar{Q}_j}}$$

$$w_j = \frac{n_{1j} n_{2j}}{n_{1j} + n_{2j}}$$

$$\bar{P}_j = \frac{n_{1j} P_{1j} + n_{2j} P_{2j}}{n_j} \quad \bar{Q}_j = 1 - \bar{P}_j$$

H_0 altında, C_0 ; asimptotik olarak $N(0,1)$ değişkeni gibi dağılır. Buna göre; α düzeyli test için kritik alan; $C_0 > Z_\alpha$ 'dır.

H_a altında en az bir j için $\pi_{1j} > \pi_{2j}$;

$$C = \frac{\sum_{j=1}^J w_j \left\{ (P_{1j} - P_{2j}) - (\pi_{1j} - \pi_{2j}) \right\}}{\sqrt{\sum_{j=1}^J w_j^2 \left\{ \frac{\pi_{1j} (1 - \pi_{1j})}{n_{1j}} + \frac{\pi_{2j} (1 - \pi_{2j})}{n_{2j}} \right\}}}$$

değişkeni $N(0,1)$ olarak asimptotik dağılır.

$\theta = \theta^* (> 1)$ için $(1 - \beta)$ 'da güç kontrolü istenirse n örneklem genişliği

$$P(C_0 > Z_\alpha \mid \theta^*) = 1 - \beta \quad (3.7)$$

olacak şekilde seçilir.

π_{1j}, π_{2j} (her j için);

$$\pi_{1j} = \frac{\pi_{2j}\theta^*}{1 - \pi_{2j} + \pi_{2j}\theta^*}$$

eşitliği ile θ^* ile ilişkilidir.

Genellikle; belirlenen $\{\pi_{ij}\}$ değerinde, belirli θ^* değeri ile birleştirilen $\{\pi_{2j}\}$ ile ilgili önsel bilgi vardır. Bu nedenle $\{\pi_{1j}\}$ değer grubu, C_0 değişkeninin asimptotik normal dağılımı ile birlikte; (3.7) koşulu,

$$1 - \Phi \left(\frac{Z_\alpha \sqrt{\sum_{j=1}^J w_j \bar{p}_j \bar{q}_j - \sum_{j=1}^J w_j (\pi_{1j} - \pi_{2j})}}{\sqrt{\sum_{j=1}^J w_j^2 \left\{ \frac{\pi_{1j}(1 - \pi_{1j})}{n_{1j}} + \frac{\pi_{2j}(1 - \pi_{2j})}{n_{2j}} \right\}}} \right) = 1 - \beta \quad (3.8)$$

$\bar{p}_j$ ve $\bar{q}_j$, gözlenmeyen rastlantı değişkenleri ise; sırasıyla, $\bar{\pi}_j = s_j \pi_{1j} + (1 - s_j) \pi_{2j}$ ve $1 - \bar{\pi}_j$ beklentileri ile $\bar{p}_j$ ve $\bar{q}_j$ yer değiştirir. Buna göre (3.8) koşulu;

$$\frac{Z_\alpha \sqrt{\sum_{j=1}^J w_j \bar{p}_j \bar{q}_j - \sum_{j=1}^J w_j (\pi_{1j} - \pi_{2j})}}{\sqrt{\sum_{j=1}^J w_j^2 \left\{ \frac{\pi_{1j}(1 - \pi_{1j})}{n_{1j}} + \frac{\pi_{2j}(1 - \pi_{2j})}{n_{2j}} \right\}}} = -Z_\beta$$

eşitliğini belirtir.

Önceden seçilmiş tasarı parametreleri $\{s_j\}$ ve $\{t_j\}$ için bu eşitlik örneklem genişliğine göre çözülür.

$$n^* = \frac{\left[Z_\alpha \sqrt{\sum_{j=1}^J t_j s_j (1-s_j) \bar{\pi}_j} + Z_\beta \sqrt{\sum_{j=1}^J t_j s_j (1-s_j) \left\{ (1-s_j) \pi_{1j} (1-\pi_{1j}) + s_j \pi_{2j} (1-\pi_{2j}) \right\}} \right]^2}{\left[\sum_{j=1}^J t_j s_j (1-s_j) (\pi_{1j} - \pi_{2j}) \right]^2}$$

Gereken örneklem genişliği; $\lceil n^* \rceil + 1$ 'e eşittir.

Olgu–Kontrol çalışmalarında, birden fazla risk faktörüyle örneklem genişliği belirtimi Rao(1986) tarafından bulunmuştur.

3.1.5.İki Aşamalı Tasarım

Klinik denemelerde veri, genellikle çalışma tamamlanmadan önce en az bir kez gözden geçirilir. Bu gözden geçirme sırasında, çalışmanın bitmesi veya devam etmesi için bir karara varılır. Bu süreç, tasarlanan iki aşamalı test metodlarının aşağıdaki türlerine neden olur:

1. Bir aşamada belirli bir karara, kesin ve açık olarak belirtilen sıfır hipotezini devam ettirmek veya iptal etmek için varılır. Yetersiz belirleyiciler nedeniyle karar verilemiyorsa aşama ikiye devam edilir.
2. Bir aşamada H_0 'a devam edip çalışmayı bitirme veya H_0 'ı iptal edip karar vermek için belirleyici arttırma, aşama ikiye devam etme.

Ayrılmış bağımlı değişken kullanılarak aynı iki aşama metodu son zamanlarda Case v.d.(1987) ve Thall v.d.(1988) tarafından çalışılmıştır. En iyi tedavi seçimi ve birkaç tedavinin etkililik testi için iki aşamalı yöntem, Thall v.d.(1988) tarafından incelenmiştir.

3.2.Nicel Bağımlı Değişkenlerle Deneyler

Bu bölümde, bağımlıların nicel değişken olduğu karşılaştırmalı deneyler incelenecektir. Bu tartışma, biyoeşdeğerlik problemini ve yaşam dağılımlarını karşılaştırmak için klinik denemeler tasarısını içerir.

3.2.1.Biyoeşdeğerliğe Göre Sınama

μ_E ve μ_S ; bir ilacın standart ve deneysel formülasyonu için biyolojik aktivite ölçüsünün ortalama değerleri olsun. Belirli A ve B sabitlerine göre;

$$A < \mu_E - \mu_S < B$$

ise, iki formülasyon biyoeşdeğerlik hipotezinin testi için kullanılacak veriye göre tasarlanmış olur. Bu problem,

$$H_0 : \mu_E - \mu_S \leq A$$

veya

$$H_0 : \mu_E - \mu_S \geq B$$

sıfır hipotezi testi için formülendir(belirtilir).

Matematiksel problemin çözümü; uygun $(1 - \beta)$ güce sahip α düzeyli testin yorumlanması için n 'in belirlenmesidir. Bağımlı değişkenin σ^2 varyanslı normal dağılımlı olduğu varsayılır.

$\bar{X}_E$ ve $\bar{X}_S$, örnekleme ortalamaları olsun. Anderson ve Hauck(1983) tarafından ileri sürülen test istatistiği aşağıdaki gibidir:

$$T = \frac{\bar{X}_E - \bar{X}_S - \frac{1}{2}(A + B)}{S_p \sqrt{\frac{2}{n}}}$$

S_p^2 ; $\nu = 2(n - 1)$ serbestlik derecesine dayanan σ^2 'nin birleştirilmiş tahmin edicisidir. T istatistiği; ν serbestlik dereceli merkezi olmayan t dağılımına sahiptir. Δ , merkezi olmama parametresi aşağıdaki formülle bulunur:

$$\Delta = \left\{ \mu_E - \mu_S - \frac{1}{2}(A + B) \right\} / \left\{ \sigma \sqrt{\frac{2}{n}} \right\}$$

Testin kritik alanı; $|T| < C$ 'dir. C;

$$P(|T| < C \mid \mu_E - \mu_S = A) = P(|T| < C \mid \mu_E - \mu_S = B) = \alpha$$

olacak şekilde belirlenir. Testin α genişliğinde olması için C'yi belirlemek zor olduğundan P değerinin hesaplanması önerilir. Böylece eşitlik için uygun sonuçlar elde edilir. $\mu_E - \mu_S = B$ eşitliğinde P değeri için tahmin aşağıdaki gibi hesaplanır:

$$p = F_v(|T_{obs}| - d) - F_v(-|T_{obs}| - d)$$

$$d = \frac{B - A}{2S_p \sqrt{\frac{2}{n}}}$$

$F_v(\cdot)$; T_v dağılımının dağılım fonksiyonudur. $p \leq \alpha$ ise formüller biyoeşdeğerliğe göre ifade edilir.

$\mu_E - \mu_S = (A + B)/2$ olduğunda gücü hesaplamak oldukça kolaydır. Bu durumda $(1 - \beta)$ 'ya eşitlenir ve ortak örneklem genişliği belirlenir. Diğer bir ifadeyle;

$$P\left(T < C \mid \mu_E - \mu_S = \frac{A + B}{2}, \sigma\right) = 1 - \beta$$

olursa n belirlenir.

$C = t_{\beta/2}(\nu)$, T_v dağılımının üst $100(\beta/2)$ yüzdeliği olduğunda bu eşitlik sağlanır. C;

$$P(T \leq C \mid \mu_E - \mu_S = B, \sigma) - P(T \leq -C \mid \mu_E - \mu_S = B, \sigma) = \alpha \quad (3.9)$$

koşulunu da sağlamalıdır.

$C = t_{\beta/2}(\nu)$ kullanılarak (3.9) eşitliği;

$$P(T(\nu, \Delta_1) \leq t_{\beta/2}(\nu)) - P(T(\nu, \Delta_1) \leq -t_{\beta/2}(\nu)) = \alpha \quad (3.10)$$

şeklinde tekrar yazılabilir. Δ_1 eşitliği aşağıdaki gibi hesaplanır:

$$\Delta_1 = \frac{(B - A)}{2} / \left\{ \sigma \sqrt{\frac{2}{n}} \right\}, \nu = 2(n - 1)$$

Bu eşitlik $\Delta_1 = \bar{\Delta}_1$ 'e göre tekrar çözülür. Bu durumda n;

$$\bar{n} = \frac{\delta \bar{\Delta}_1^2 \bar{\sigma}^2}{(B-A)^2} \quad (3.11)$$

şeklinde elde edilir.

$\bar{\sigma}^2$; σ^2 'nin önsel tahminidir.

Merkezi dağılım ile merkezi olmayan t dağılımı tahmin edilerek ve sol taraftaki ikinci terim ihmal edilerek (3.10) eşitliği;

$$P \left(T(\nu) \leq t_{\beta/2}(\nu) - \frac{(B-A)}{2\sigma\sqrt{\frac{2}{n}}} \right) = \alpha$$

şekline indirgenir.

$$n = \frac{8(t_{\beta/2}(\nu) + t_\alpha(\nu))^2 \sigma^2}{(B-A)^2}, \nu = 2(n-1)$$

olduğunda bu eşitlik sağlanır. n^* kesin sonucunu bulmak için bu eşitlik tekrar çözülebilir.

Teorem 3.2: $\mu_E - \mu_S = (A+B)/2$ eşitliğinde $(1-\beta)$ güç elde edilerek $H_a: A < \mu_E - \mu_S < B$ alternatifine karşı $H_0: \mu_E - \mu_S \leq A$ veya $H_0: \mu_E - \mu_S \geq B$ sıfır hipotezini α düzeyinde test etmek için gereken ortak örneklem genişliği tahmini; $[n^*] + 1$ 'dir. n^* eşitliği, aşağıda verilmiştir:

$$n^* = \frac{8(t_{\beta/2}(\nu) + t_\alpha(\nu))^2 \sigma^2}{(B-A)^2}, \nu = 2(n-1)$$

Bu eşitlikteki $t_{\beta/2}(\nu) + t_\alpha(\nu)$ ifadesinin (3.11) eşitliğinde elde edilen $\bar{\Delta}_1$ için tahmin olduğuna dikkat edilmelidir.

3.2.2. İki Süreç, İki Tedavi, Çapraz Çalışması

Klinik denemelerde çoğunlukla kullanılan deneysel ortam, çapraz tasarımıdır. 0-1 bağımlı değişkenli bu çalışmadan Bölüm 3.1.1’de kısaca bahsedilmişti. Sürekli bağımlı değişken ile bu modelin detaylı tartışmasına bu bölümde girilecektir.

A ve B, rastgele olarak $2n$ seçilen deneklere uygulanan iki tedavi olsun. $2n$ deneklerden her bir n iki gruba rastgele ayrılacaktır. B ile izlenen A tedavi etkisi, birinci grubun her deneği için test edilecektir. A ile izlenen B etkisi, ikinci grubun her deneği için test edilecektir. Bu durumda deneysel düzen aşağıdaki gibi oluşur:

Denekler								
Periyod	1	2	...	n	n+1	n+2	...	2n
I	A	A	...	A	B	B	...	B
II	B	B	...	B	A	A	...	A

3.4. Birbirini İzleyen Tedavilerde Deneysel Düzen

Uygun model aşağıdaki gibidir:

$$Y_{1j1} = \mu + \rho_1 + \gamma_{1j} + \tau_A + e_{1j1}$$

$$Y_{2j1} = \mu + \rho_1 + \gamma_{2j} + \tau_B + e_{2j1}$$

$$Y_{1j2} = \mu + \rho_2 + \gamma_{1j} + \tau_B + \delta_A + e_{1j2}$$

$$Y_{2j2} = \mu + \rho_2 + \gamma_{2j} + \tau_A + \delta_B + e_{2j2}, \quad j = 1, 2, \dots, n$$

Bu model, $H_0: \tau_A = \tau_B$ hipotezini α önemlilik düzeyindeki testte güç kontrolüyle ortak örneklem genişliğini belirlemek için oluşturulur. Bunun için $\delta_A = \delta_B$ yani her iki tedavinin artık etkilerinin aynı olduğu varsayılır.

$$Z_{ij} = Y_{ij1} - Y_{ij2}; \quad i = 1, 2; \quad j = 1, 2, \dots, n$$

olsun. $H_0: \tau_A = \tau_B$ hipotezinin A testi, $\{Z_{1j}\}$ ve $\{Z_{2j}\}$ iki bağımsız örneklem ilişkisinde iki örneklem problemine eşit olur.

Dağılımın varsayımları aşağıda verilmiştir:

- i) e_{ijk} ; bağımsız ve sıfır ortalama σ^2 varyanslı normal dağılıma sahiptir.

- ii) γ_{ij} ; bağımsız ve sıfır ortalamalı σ_b^2 varyanslı normal dağılıma sahiptir.
- iii) e_{ijk} ve γ_{ij} , bağımsız olarak dağılır.

$H_0 : \tau_A = \tau_B$ hipotezi iki örneklem t testi ile test edilir. “ $\sigma_1^2 = \sigma_2^2$ Olduğunda ve Ortak Varyans Bilinmediğinde μ_1 ve μ_2 İle İlgili Hipotez Testi” başlığındaki anlatıma benzer bir tartışma aşağıda elde edilir:

Teorem 3.3: İki yanlı, α önemlilik düzeyinde $H_0 : \tau_A = \tau_B$ sıfır hipotezi test edilirken $|\tau_A - \tau_B| = (> 0)$ olduğunda $(1 - \beta)$ güç elde etmek için gereken ortak örneklem genişliği $n; \left(\frac{\nu}{2} + 1\right)$ ’e eşittir. ν ;

$$n = \left(\frac{\nu}{2} + 1\right) = \frac{2(2\sigma^2)(t_{\alpha/2}(\nu) + t_{\beta}(\nu))^2}{\Delta^2}$$

eşitliğini sağlayan pozitif en küçük tam sayıdır.

Tek yanlı test için $\alpha/2$ yerine α koyulur.

e_{ijk} ve γ_{ij} için normallik varsayımının doğru olmadığı varsayılın. Bu durumda doğrudan tedavi etkilerinin eşitliği, Wilcoxon iki örneklem testi ile test edilebilir. Bu durumda ortak örneklem genişliği tahmini Noether(1987)’in aşağıdaki çalışması ile belirlenebilir.

3.2.3.Yaşam Dağılımlarına Benzerliğine Göre Klinik Denemeler

Klinik denemeler, aşağıdaki formata sahip iki tedavi altında yaşam karşılaştırılması için incelenir. Çalışmanın başlamasından sonra hastaların denemeye girmeleri ve rastgele olarak tedavilere ayrılmaları süresinde bir iyileşme periyodu vardır. İyileşme periyodu; hastaların gözlem altında olacakları sürede “takip (izleme) periyodu” olarak adlandırılan bir periyod ile izlenir. Bu çalışmalarda, bazı hastalar için ölüm (ölüm zamanları olarak bakılacaktır) gibi bazı kritik olaylara göre zamanlar gözlemlenemez ve sansürlü olarak söylenir. Sansürleme genelde iki şekilde olur:

- 1) Hastalar iyileşme ve takip dönemleri boyunca, takip için geri kalmış veya çalışmadan ayrılmış olabilirler.
- 2) Hastalar, çalışmaya son verildiğinde hâlâ hayatta (yaşıyor) olabilirler.

Bunun gibi klinik denemeler tasarımı, α düzeyli test için istenen gücü elde etmede gereken deneme süresinin belirlenmesinden oluşur. Burada; yaş dağılımları, iyileşme oranı, takip dönemi ve takip için kayıp oranı belirlidir. Böylece, örneklem genişliği bulma problemi, elde edilen koşullarla deneme süresinin belirlenmesine indirgenir.

Aşağıdaki tartışma, Rubinstein v.d.(1981) çalışmasına dayandırılır. Önceden bahsedilen tasarım problemi, aşağıdaki varsayımların uygulanmasıyla çözülür:

- i) T yıllık iyileşme periyodu boyunca hastalar, her yıl için “a” oranlı Poisson sürecine göre denemeye girer. Bu, kontrol veya deneysel tedaviye alınan her hasta için yaklaşık olarak eşittir. Takip periyodu, eklenen τ yıl(lar)ıdır.
- ii) Ölüm zamanı, her bir tedavi için ortak üstel dağılımlı ve bağımsızdır. Kontrol ve deneysel tedaviler için anlık ölüm olasılığı (hazard rate), sırasıyla λ_c ve λ_e ’dir.
- iii) Zaman, “takipe göre kayıp” için girdiden bağımsızdır ve onların hem giriş hem de ölüm zamanları bağımsızdır. Bu zamanlar; deneysel grup için Φ_e , kontrol grubu için Φ_c oranlı üstel dağılıma uyarlar.

Test edilecek sıfır hipotezi; $H_a : \lambda_c > \lambda_e$ alternatifine karşı $H_0 : \lambda_c = \lambda_e$ ’dir.

İyileşme periyodu süresi T, $\lambda_c = r^* \lambda_e$ ($r^* > 1$) alternatifi, $(1 - \beta)$ güce sahip olan $H_0 : \lambda_c = \lambda_e$ hipotezinin testi için belirlenir.

D_c ve D_e ; iki tedavi grubundaki ölümlerin sayısı olsun. Ayrıca; T_c ve T_e ; sırasıyla kontrol ve deneysel gruplar için riskteki toplam zaman (yaşam zamanları ve sansürlü zamanların toplamı) olsun. Test; $\hat{\lambda}_c = \frac{D_c}{T_c}$, $\hat{\lambda}_e = \frac{D_e}{T_e}$ tahmin edicilerine dayanır.

$\ln \hat{\lambda}_i$ ’nin tahmini dağılımı, $N(\ln \lambda_i, 1/D_i)$, $i = c, e$ ’dir. Asimptotik dağılım kullanılmasıyla, testin kritik alanı aşağıdaki gibi olur:

$$\ln \left(\frac{\hat{\lambda}_c}{\hat{\lambda}_e} \right) > Z_\alpha \sqrt{\frac{1}{D_c} + \frac{1}{D_e}}$$

Bahsedilen güç koşulundan sonra $H_a : \lambda_c > \lambda_e$ eşitliği;

$$1 - \Phi \left(Z_\alpha - \frac{\ln r^*}{\sqrt{\frac{1}{D_c} + \frac{1}{D_e}}} \right) = 1 - \beta$$

olarak tekrar yazılır.

Buradan;

$$\frac{(\ln r^*)^2}{(Z_\alpha + Z_\beta)^2} = \frac{1}{D_c} + \frac{1}{D_e} \quad (3.12)$$

şeklinde düzenlenebilir.

Düzenlenen çalışmada D_c ve D_e , gözlenmemiş rastlantı değişkenleridir. (3.12) eşitliğinde bunların beklenen değerleri D_c ve D_e yerine koyularak;

$$\frac{(\ln r^*)^2}{(Z_\alpha + Z_\beta)^2} = \frac{1}{E(D_c)} + \frac{1}{E(D_e)} \quad (3.13)$$

bulunur.

$$E(D_i) = \frac{\alpha T \lambda_i}{2 \lambda_i^*} \left\{ 1 - \frac{e^{-\lambda_i T}}{\lambda_i^* T} (1 - e^{\lambda_i T}) \right\} \quad (3.14)$$

$$\lambda_i^* = \lambda_i + \Phi_i, i = c, e$$

olduğu görülür.

(3.13) eşitliğindeki $E(D_i)$ için (3.14) ifadesi kullanılarak bulunan eşitlik T'ye göre çözülür. $\Phi_i = 0$, $\Phi_i = \lambda_i / 4$ için $T + \tau$ 'ya göre tablo, Rubinstein v.d.(1981) tarafından hazırlanmıştır.

3.3.Lojistik Regresyon İçin Örneklem Genişliği ve Güç

3.3.1.İkili Oran Karşılaştırması İçin Örneklem Genişliği

X ikili olduğunda, çalışma çoğu kez ikili bağımlılık üzerine (yani X'in iki düzeyi) iki grup karşılaştırmasına bakar. Gruplar için π_1 ve π_2 başarı olasılıklarının özdeş olduğu hipotez testi için, bağımlılığa göre çapraz sınıflama yapılan 2x2lik tabloya göre Ki-Kare testi yapılabilir. P değeri sabit α önemlilik düzeyinden küçükse sıfır hipotezi reddedilir. Örneklem genişliğini belirlemek için gerçekte önemli olarak kabul edilen π_1 ve π_2 arasındaki farkın sabit genişliğine göre β hata olasılığını belirlemek gerekir. Bu genişliğin sonucuna göre β ; α düzeyinde " H_0 sıfır hipotezini reddetme" hata olasılığıdır. Bu durumda α ; 1. tip hata olasılığı ve β ; 2. tip hata olasılığıdır. Testin gücü; $(1 - \beta)$ 'ya eşittir.

Her grupta eşit örneklem genişliği kullanan bir çalışmaya, aşağıdaki formülle genişlik tahmini yapılır:

$$N_1 = N_2 = \frac{(Z_{\alpha/2} + Z_{\beta})^2 [\pi_1(1 - \pi_1) + \pi_2(1 - \pi_2)]}{(\pi_1 - \pi_2)^2}$$

Her grupta hata olasılıklarını tahmin etmek gerekir. Bu formül için π_1 , π_2 , α ve β değerleri gerekir.

Bu formül, $(\pi_1 - \pi_2)$ 'ye göre güven aralığı için de gereken örneklem genişliği tahmini yapar. Bu durumda α ; aralık için hata olasılığı, β ; sonucun makul eksikliğini gösteren güven aralığı için hatadır. Yani, bu, sıfır değerini içerir.

Bu örneklem genişliği formülü yaklaşık sonuç verir. Gereken gerçek değerin çok az altında tahmin yapar. π_1 ve π_2 için kaba bir tahmin yapan normallığe göre ise pek çok pratik çalışma için yeterlidir. Fleiss daha kesin sonuç veren bir formül geliştirmiştir.

2x2lik tabloda iki oran karşılaştırması testi için hipotez testi,

$$\text{lojit}(\pi) = \alpha^* + \beta^* x$$

şeklindeki lojistik regresyon modelindeki parametrelerden birine eşittir. (1. ve 2. tip hata olasılıklarıyla parametrelerin karışmaması için * işareti kullanılmıştır.) 1. grup için

$X = 1$ ve 2. grup için $X = 0$ atarsak, başarı olasılığının lojiti 1. grup için $\alpha^* + \beta^*$ 'a ve 2. grup için α^* 'a eşit olur. Özdeş olasılıklar, özdeş lojitelere veya özdeş oddslara veya $\beta^* = 0$ 'a karşılık gelir. Böylece bu örnek, basit lojistik regresyon modeline göre örneklem genişliği belirlemekle ilişkili olur. (Agresti, 1996)

3.3.2. Nicel (Sayısal) Tahmin Edicilerle Örneklem Genişliği

Nitel göstergelerden çok nicel olan X ile

$$\text{lojit}(\pi) = \alpha^* + \beta^* x$$

şeklindeki modele göre, X değerinin dağılımına bağlı $H_0: \beta^* = 0$ testinde belirli güç elde etmek için örneklem genişliği gereklidir. Ortalama X 'te $\bar{\pi}$ başarı olasılığını tahmin edilmelidir. Sonuç genişliği, ortalama X 'e bakan başarı olasılığının aynı standart sapması için o aşamaya gelince başarı olasılıklarını karşılaştıran θ odds oranıdır.

$\lambda = \log(\theta)$ olsun. F.Y.Hsieh(1989); tek yanlı teste göre aşağıdaki yaklaşık örneklem genişliği formülünü geliştirdi:

$$N = \frac{[Z_\alpha + Z_\beta \exp(-\lambda^2 / 4)]^2 (1 + 2\bar{\pi}\delta)}{\pi\lambda^2} \quad (3.15)$$

$$\delta = \frac{1 + (1 + \lambda^2) \exp(5\lambda^2 / 4)}{1 + \exp(-\lambda^2 / 4)}$$

N değeri, $\bar{\pi}$; 0.5'e yakın bulunduğu ve $|\lambda|$; sıfır değeri 0'dan uzak bulunduğu azalır (küçülür). Yaklaşık olarak, normal dağılımlı açıklayıcı değişkenin türetimi varsayılır.

3.3.3. Çoklu Lojistik Regresyonda Örneklem Genişliği

Bir çoklu lojistik regresyon modeline kısmi etkileri (sonuç) fark etmek için büyük örneklem genişliği gerekir. R ; ilgilenilen tahmin edici X ile modeldeki diğerleri arasındaki çoklu korelasyonu gösterebilir. N 'i bulma için (3.15) formülü $(1 - R^2)$ 'ye bölünür. Θ formülde, $\bar{\pi}$; bütün açıklayıcı değişkenlerin ortalama değerindeki olasılığı belirtir ve odds oranı, diğerlerinin ortalama düzeylerindeki ilgilenilen tahmin edicinin etkisini gösterir.

Bu formüller sadece örneklem genişliğinin kabataslak (yaklaşık) belirtimini bulur. Pek çok uygulamada, $\bar{\pi}$ ve R olmak üzere göstergeler için sadece kaba bir tahmin vardır ve açıklayıcı değişken normal dağılımdan uzak olabilir.(Agresti, 1996)



4. UYGULAMA

Bu bölümde, ikinci ve üçüncü bölümlerde bahsedilen örneklem genişliği belirlemede kullanılan yöntemlerden bazılarının uygulamaları verilecektir. Uygulamalar, Win Episcpe 2.0, Minitab, S-Plus istatistik paket programlarında yapılmıştır.

4.1. Anakütle Ortalamasının Tahmininde Örneklem Genişliği Belirleme

Bir sağlık ocağı bölgesinde yaşayan 15-44 yaşında evli ve en az bir kez doğum yapmış 200 kadının bulaşıcı hastalıklar bilgi puanının 54 ortalama ve 18 standart sapma ile normal dağıldıkları bilinmektedir. Mutlak hatanın 16'dan büyük olmaması istendiğine göre %95 güvenirlilikle gereken örneklem genişliği tahmini aşağıda verilmiştir:

Win Episcpe 2.0:

Input of DATA:		RESULTS:	
Population Size:	200	Required sample size:	
	Unknown	Exact calculation:	5
Expected Standard Deviation:	18	Approximated calculation:	5
Expected Absolute Error:	16		
Level of Confidence (%):	95 %		
		Calculate	Close

4.1. Anakütle Ortalamasının Tahmininde Örneklem Genişliği Belirleme

Win Episcpe 2.0 Çözümü

Minitab:

06.01.2005 12:35:00

Power and Sample Size

1-Sample Z Test

Testing mean = null (versus > null)
Calculating power for mean = null + difference
Alpha = 0,05 Sigma = 18

Difference	Sample Size	Target Power	Actual Power
16	8	0,8000	0,8077

Specify values for any two of the following:

Sample sizes:

Differences:

Power values:

Sigma:

Options... Help OK Cancel

4.2. Anakütle Ortalamasının Tahmininde Örneklem Genişliği Belirleme**Minitab Çözümü****S-Plus 2000:**

*** Power Table ***

mean.null	sdi	mean.alt	delta	alpha	power	n1
54	18	80	26	0.05	0.8	4

Model Options Printout

Select Compute: ☒ Sample Size ☐ Power ☐ Min. Difference

Sample Type:

Probabilities: Alpha(s): Power(s):

Sample Sizes: N1: N2: N2/N1:

Standard Deviations: Sigma1: Sigma2: Sigma(X2 - X1):

Null Hypothesis: Mean: Mean1:

Alternative Hypothesis: Alt Mean: Mean2: Test Type:

Results: Save As: ☒ Print Results:

OK Cancel Apply > current Help

4.3. Anakütle Ortalamasının Tahmininde Örneklem Genişliği Belirleme**S-Plus 2000 Çözümü**

4.2.Ortalamalar Arasındaki Farklılığın Karşılaştırılmasında Örneklem Genişliği Belirleme

Akut miyokard enfarktüsü geçirmiş hastalarda ve normal kontrollerde serumda selenyum (se) düzeyi ölçülmektedir. Bu iki gruptaki se ortalamalarının karşılaştırmak istenildiği bir çalışma için, akut miyokard enfarktüsü geçirmiş hastalarda ortalama selenyum 60 ve normal grupta 74 ve standart sapma 15 olarak alınırsa selenyum örnek büyüklüğünü tahmin ediniz.

Win Episcpe 2.0:

Input of DATA:	
Expected mean in group 1:	60
Expected mean in group 2:	74
Expected Standard Deviation:	15
Level of Confidence (%):	95 %
Power (%):	85 %

RESULTS:	
Difference between means:	14.000
Sample Size	One-Tailed Two-Tailed
Independent samples:	19 23
Paired samples:	12 14

Buttons: Calculate, Close

4.4.Ortalamalar Arasındaki Farklılığın Karşılaştırılmasında Örneklem Genişliği Belirleme

Win Episcpe 2.0 Çözümü

Minitab:

Session

06.01.2005 12:35:00

Power and Sample Size

2-Sample t Test

Testing mean 1 = mean 2 (versus not =)
Calculating power for mean 1 = mean 2 + difference
Alpha = 0,05 Sigma = 15

Difference	Sample Size	Target Power	Actual Power
14	22	0,8500	0,8563

Power and Sample Size for 2-Sample t

Specify values for any two of the following:

Sample sizes:

Differences: 14

Power values: 0,85

Sigma: 15

Options...

Help OK Cancel

Worksheet

	C1	C2	C3	C4	C5
1					

4.5.Ortalamalar Arasındaki Farklılığın Karşılaştırılmasında Örneklem Genişliği Belirleme**Minitab Çözümü****S-Plus 2000:**

Report

*** Power Table ***

mean1	sd1	mean2	sd2	delta	alpha	power	n1	n2
60	15	74	15	14	0.05	0.85	21	21

Model Options Plotout

Select

Computer: ☒ Sample Size ☐ Power ☐ Min. Difference

Sample Type: Two Sample

Probabilities

Alpha(s): 0.05

Power(s): 0.85

Sample Sizes

N1:

N2:

N2/N1: 1

Standard Deviations

Sigma1: 15

Sigma2: Sigma1

Sigma*(X2 - X1):

Null Hypothesis

Mean:

Mean1: 60

Alternative Hypothesis

Alt Mean:

Mean2: 74

Test Type: two.sided

Results

Save As:

☒ Print Results

OK Cancel Apply K current Help

4.6.Ortalamalar Arasındaki Farklılığın Karşılaştırılmasında Örneklem Genişliği Belirleme**S-Plus 2000 Çözümü**

4.3. Anakütle Oranının Tahmininde Örneklem Genişliği Belirleme

Bir bölgede yaşayan 5yaş altı 1000 çocukta tüberküloz oranı saptanmak istiyor. Gerçek oranın %25'i aşmayacağı tahmin edildiğinde, %90 güven düzeyinde, bu oranın ± 7 sınırları içinde kalması için örneğe kaç çocuk seçilmesi gerekir?

Win Episcpe 2.0:

% Expected Prevalence	% Level of Confidence				
	90	95	97.5	99	99.5
0	1	1	1	1	1
10	50	71	93	122	145
20	89	126	165	217	258
30	116	165	216	285	338
40	133	189	247	325	386
50	139	196	257	339	403
60	133	189	247	325	386
70	116	165	216	285	338
80	89	126	165	217	258
90	50	71	93	122	145
100	1	1	1	1	1

4.7. Anakütle Oranının Tahmininde Örneklem Genişliği Belirleme

Win Episcpe 2.0 Çözümü

Minitab:

Session

06.01.2005 12:35:00

Welcome to Minitab, press F1 for help.

Power and Sample Size

Test for One Proportion

Testing proportion = 0,5 (versus not = 0,5)
Alpha = 0,05

Alternative Proportion	Sample Size	Target Power	Actual Power
0,250000	29	0,8000	0,8012

Specify values for any two of the following:

Sample sizes:

Alternative values of p:

Power values:

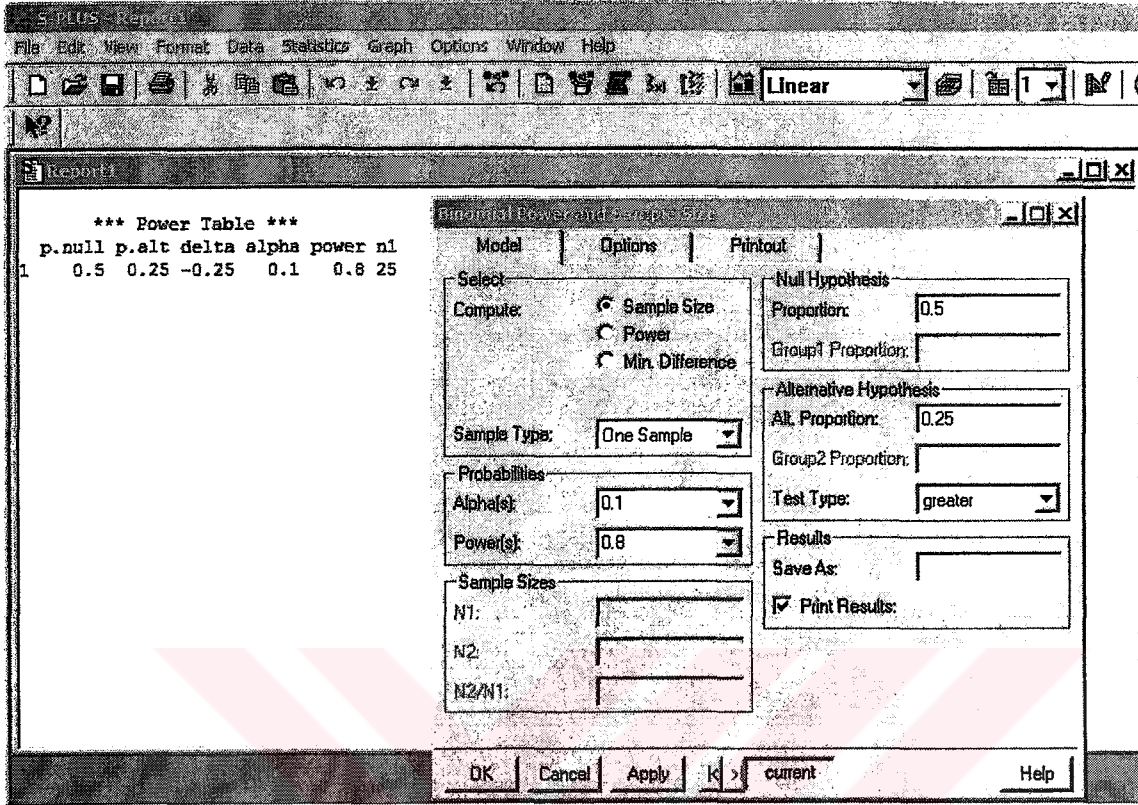
Hypothesized p:

Options...

Worksheet1

	C1	C2	C3	C4	C5
1					
2					
3					

4.8. Anakütle Oranının Tahmininde Örneklem Genişliği Belirleme**Minitab Çözümü**

S-Plus 2000:**4.9. Anakütle Oranının Tahmininde Örneklem Genişliği Belirleme****S-Plus 2000 Çözümü**

4.4.Oranlar Arasındaki Farklılığın Karşılaştırılmasında Örneklem Genişliği Belirleme

Aynı hastalığın tedavisi için uygulanan iki cerrahi yöntemden birisinde % 8, diğerinde ise % 17 oranında komplikasyon görüldüğü sanılmaktadır. İkinci cerrahi yöntemin birinciye göre önemli ölçüde fazla komplikasyona neden olduğunu %1 güven düzeyinde ve %85 test gücü ile söyleyebilmek için kaçar kişilik gruplar üzerinde çalışmak gerektiğini tahmin ediniz.

Win Episcpe 2.0:

Input of DATA:		RESULTS:	
Proportion in population 1 (%)	8	Sample size per treatment group:	
Proportion in population 2 (%)	17	One-Tailed:	301
Level of Confidence (%)	99 %	Two-Tailed:	347
Power (%)	85 %		
		Calculate	Close

4.10. Oranlar Arasındaki Farklılığın Karşılaştırılmasında Örneklem Genişliği Belirleme

Win Episcpe 2.0 Çözümü

Minitab:

06.01.2005 12:35:00

Power and Sample Size

Test for Two Proportions

Testing proportion 1 = proportion 2 (versus <)
Calculating power for proportion 2 = 0,17
Alpha = 0,01

Proportion 1	Sample Size	Target Power	Actual Power
0,080000	304	0,8500	0,8505

Specify values for any two of the following:

Sample sizes: _____

Proportion 1 values: 0,08

Power values: 0,85

Proportion 2: 0,17

Options...

Help OK Cancel

4.11. Oranlar Arasındaki Farklılığın Karşılaştırılmasında Örneklem Genişliği Belirleme
Minitab Çözümü

S-Plus 2000:

*** Power Table ***

	p1	p2	delta	alpha	power	n1	n2
1	0.08	0	-0.08	0.01	0.85	159	159
2	0.08	0.17	0.09	0.01	0.85	NA	NA

Model Options Printout

Select

Compute: ☒ Sample Size ☐ Power ☐ Min. Difference

Sample Type: Two Sample

Probabilities

Alpha(s): 0.01

Power(s): 0.85

Sample Sizes

N1: _____

N2: _____

N2/N1: 1

Null Hypothesis

Proportion: _____

Group1 Proportion: 0.0

Alternative Hypothesis

Alt. Proportion: _____

Group2 Proportion: 0.17

Test Type: less

Results

Save As: _____

☒ Print Results

OK Cancel Apply K current Help

4.12. Oranlar Arasındaki Farklılığın Karşılaştırılmasında Örneklem Genişliği Belirleme
S-Plus 2000 Çözümü

4.5. Bir Anakütlerde Hastalığı Teşhis Etmek İçin Gereken Örneklem Genişliğinin Belirlenmesi

Göğüs kanseri olan 100 kadından 65'inde pozitif test sonucu alınmıştır. Teşhisin doğru olup olmadığını araştırmak için %95 güvenilirlikle gereken örneklem genişliğini tahmin ediniz.

Win Episcopes 2.0:

N° Diseased	% Diseased	Sample Size	N° Diseased	% Diseased	Sample Size
1	1.00	95.00	11	11.00	22.65
2	2.00	77.25	12	12.00	20.88
3	3.00	62.53	13	13.00	19.35
4	4.00	51.92	14	14.00	18.01
5	5.00	44.12	15	15.00	16.84
6	6.00	38.32	16	16.00	15.78
7	7.00	33.77	17	17.00	14.86
8	8.00	30.14	18	18.00	14.03
9	9.00	27.18	19	19.00	13.27
10	10.00	24.72	20	20.00	12.59

4.13. Bir Anakütlerde Hastalığı Teşhis Etmek İçin Gereken Örneklem Genişliğinin Belirlenmesi

Win Episcopes 2.0 Çözümü

4.6.Eşleştirilmemiş Olgu–Kontrol Çalışmalarında Örneklem Genişliği Belirleme

Reye sendromu olan hastalarda Aspirin kullanımıyla ilgili olgu-kontrol çalışması yapılıyor. Reye olmayan ama aspirin kullananların oranı %36'dır. Viral hastalıktan önce aspirin kullanımının Reye sendromuna yakalanma olasılığını 8 kat arttırdığı biliniyorsa %90 güvenilirlik ve % 80 test gücüyle gereken örneklem genişliğini tahmin ediniz.

The screenshot shows the 'Sample Size' window in WinEpiscope 2.0. The 'Input of DATA' section contains the following values: 'N° of controls per case' is 200, 'Estimated exposure rate Controls (%)' is 36, 'Minimal OR to be detected signif.' is 8, 'Level of Confidence (%)' is 90 %, and 'Power (%)' is 80 %. The 'RESULTS' section shows 'Estimated exposure rate Cases (%)' as 81.82. Below this, a table shows the 'Sample size per group' for 'Cases' and 'Controls' for both 'One-Tailed' and 'Two-Tailed' tests. The 'Calculate' button is highlighted.

	Cases	Controls
One-Tailed:	4	848
Two-Tailed:	6	1192

4.14. Eşleştirilmemiş Olgu–Kontrol Çalışmalarında Örneklem Genişliği Belirleme

Win Episcope 2.0 Çözümü

4.7.Eşleştirilmiş Olgu Kontrol Çalışmalarında Örneklem Genişliği Belirleme

BCG'nin çocukluk çağı tüberkülozunu önleme konusu araştırılıyor. Tüberkülozlu hastalarda ve kontrollerinde aşılama oranlarını karşılaştırmak için olgu-kontrol çalışması yapılmak isteniyor. Kontrollerin %35'inin aşısız olduğu tahmin ediliyor. Araştırmacı %10 önem düzeyinde ve %85 test gücüyle odds oranının 1'den anlamlı şekilde farklı olduğunu göstermek istemektedir. Odds oranının 3 olması durumunda gruplar arasındaki farkın önemli olacağı düşünülüyorsa her grupta en az kaç kişi yer almalıdır?

Input of DATA:	
Estimated exposure rate Controls (%)	35
Level of Confidence (%)	90 %
Minimal OR to be detected signif.	3
Power (%)	85 %

RESULTS:	
Estimated exposure rate Cases (%)	81.76
Sample Size:	97
Cases:	48
Controls:	48

Buttons: Calculate, Close

4.15. Eşleştirilmiş Olgu Kontrol Çalışmalarında Örneklem Genişliği Belirleme
Win Episcop 2.0 Çözümü

4.8. Kohort (Grup İzleme) Çalışmalarında Örneklem Genişliği Belirleme

Çok merkezli bir klinik araştırma ile bir kanser türünün tedavisinde kullanılan iki farklı yöntem karşılaştırılmak isteniyor. A ve B tedavileri hastalara rastgele uygulanıyor ve hastalar 5 yıl boyunca izleniyor. Kullanılmaya yeni başlanan A tedavisi, 5 yıl boyunca nüks olasılığını yarıya indirebilirse başarılı kabul edilecektir. B tedavisini alan hastaların %30'unda nüks görüldüğü biliniyor. “Tedaviler arasında fark yoktur” hipotezini %10 önem düzeyinde %95 test gücüyle reddetmek için araştırmacı her gruba en az kaçar kişi seçmelidir?

Input of DATA:	
% Non-Exposed Diseased animals:	30
Minimal RR to be detected signif.:	0,5
Power (%):	95 %
Level of Confidence (%):	90 %
RESULTS:	
% Exposed Diseased animals:	15,00
Sample Size of each Cohort:	163

Buttons: Calculate, Close

4.16. Kohort (Grup İzleme) Çalışmalarında Örneklem Genişliği Belirleme

Win Episcopy 2.0 Çözümü

4.9. Lojistik Regresyonda Örneklem Genişliği Belirleme

Orta yaşlı anakütleye göre kolesterol düzeyi ile kalp hastalığı olasılığının bağımlılığı araştırılmak isteniyor ve bu amaçla kolesterol artışı için $\beta^* > 0$ risk artışı alternatifine karşı bağımsızlık ($\beta^* = 0$) testi inceleniyor. Şiddetli kalp hastalığı olasılığının, ortalama kolesterol düzeyinde 0.08 olduğu daha önce yapılmış çalışmalardan elde ediliyor. Kolesterolde standart sapma artışı olduğunda bu olasılıkta %50 artışa (yani 0.12'ye) duyarlı olacak test isteniyor. Bu bilgilere göre örneklem genişliğini hesaplayınız.

$$odds = \frac{0.08}{0.92} = 0.087$$

$$odds = \frac{0.12}{0.88} = 0.136 \rightarrow \text{Ortalamaya bakan standart sapma odds'u}$$

$$\theta = \frac{0.136}{0.087} = 1.57 \rightarrow \text{Odds Oranı}$$

$$\lambda = \log(\theta) = 0.450$$

$$\lambda^2 = (0.450)^2 = 0.202$$

$$\alpha = 0.05 \rightarrow Z_{0.05} = 1.645$$

$$\beta = 0.10 \rightarrow Z_{0.10} = 1.28$$

$$\delta = \frac{1 + (1 + \lambda^2) \exp(5\lambda^2/4)}{1 + \exp(-\lambda^2/4)} = \frac{1 + (1.202) \exp(5(0.202)/4)}{1 + \exp(-0.202/4)} = \frac{2.548}{1.951} = 1.306$$

$$N = \frac{[Z_\alpha + Z_\beta \exp(-\lambda^2/4)]^2 (1 + 2\delta)}{\pi \lambda^2}$$

$$= \frac{[1.645 + 1.28 \exp(-0.202/4)]^2 (1 + 2(0.08)(1.306))}{(0.08)(0.202)} = 612$$

Bu örneğe göre α ve β 'nin farklı değerleri için hesaplanan örneklem genişlikleri aşağıdaki tabloda verilmiştir:

α / β	0,10	0,11	0,12	0,13	0,14	0,15	0,16	0,17	0,18	0,19	0,20
0,01	941	916	891	867	842	824	800	782	768	750	732
0,02	798	775	752	730	708	690	669	652	639	623	607
0,03	717	695	674	652	632	615	595	579	567	552	536
0,04	658	637	617	596	576	561	541	526	515	500	485
0,05	612	592	572	553	534	519	500	485	474	460	447
0,06	572	553	534	515	496	482	464	450	439	426	412
0,07	544	525	506	488	470	455	438	424	414	401	388
0,08	516	497	479	461	444	430	413	400	390	377	364
0,09	489	471	453	436	419	405	389	376	366	354	342
0,10	466	448	431	414	398	385	369	356	347	335	323

4.1.Lojistik regresyonda çeşitli α ve β değerlerine göre

örneklem genişliği hesaplamaları

Buna göre; güvenilirlik sabit tutulup güç artırıldığında örneklem genişliğinin azaldığı, güç sabit tutulup güvenilirlik artırıldığında örneklem genişliğinin yine azaldığı görülür.

SONUÇ ve ÖNERİLER

“Klinik Denemelerde Örneklem Genişliği Belirleme” başlıklı bu tezin uygulamasında Win Episcopo 2.0, Minitab ve S-Plus 2000 paket programları kullanılmıştır. Her bir paket programda elde edilen sonuçlar yaklaşık olarak eşit olmasına rağmen bazı farklılıklar göstermektedir. Farklılık; bu paket programların kullandığı örnek büyüklüğü hesaplama yaklaşımlarından kaynaklanmaktadır. Ancak uygulamacılar için bunlardan bir tanesinin kullanılması yeterli olacaktır.

Bu tezin içerisinde verilen ancak uygulamada gösterilmeyen önsel bilgilerin yani daha önceki benzer çalışma gruplarından elde ettiğimiz subjektif bilgilerin hesaplamaya katıldığı Bayesci yaklaşımın daha detaylı incelenmesi ve uygulanabilirliğinin gösterimi gelecekte yapılacak çalışmalar için bize yön verecektir.



KAYNAKLAR

- [1] Agresti, A. (1996). *An Introduction to Categorical Data Analysis*, John Wiley&Sons, Inc.
- [2] Anderson, S., and Hauck, W.W. (1983). A new procedure for testing equivalence in comparative bioavailability and other clinical trials. *Commun. in Statist. A12*, 2663-2692.
- [3] Borenstein, M., Rothstein, H., and Cohen, J.(1997), *Power and Precision*, Biostat, Teaneck, NJ. Software for MS-DOS systems.
- [4] Bowman, K. O., and Kastenbaum, M. A. (1975). Sample size requirement: single and double classification experiments. In *Selected Tables in Mathematical Statistics*, Vol. 3, 111–232 (Edited by Institute of Mathematical Statistics), American Methemathical Society, Providence, R1.
- [5] Case, L.D., Morgan, T.M., and Davies, C.E. (1987). Optimal restricted two-stage designs. *Controlled Clinical Trials* 8, 146-156.
- [6] Castelloe, J.(2000), "Sample Size Computations and Power Analysis with the SAS Systems," in *Proceedings of the Twenty-Fifth Annual SAS User's Group International Conference*, Cary, NC, SAS Institute, Inc., Paper 265-25.
- [7] Chapman, D.G. (1950). Some two-sample tests. *Ann. Math. Statist.* 21, 601–606
- [8] Cochran, W. G., and Cox, Gm M. (1957). *Experimental Designs* (Second Edition), John Wiley and Sons, New York, NY.
- [9] Cohen, J. (1987). *Statistical power Analysis* (Second Edition), Lawrence Erlbaum Associates, Inc. , Hillsdale, NJ.
- [10] Cohen, J.(1988), *Statistical Power Analysis for the Behavioral Sciences*, Academic Pres, New York, 2nd edn.
- [11] Connor, R. J. (1987). Sample Size for testing differences in proportions for the paired sample design. *Biometrics* 43, 207–212
- [12] Desu, M. M. & Raghavarao, D. (1990). *Sample size methodology*.
- [13] Dunnett, C.W., and Sobel, M. (1955). Approxmation to the probability integral and certain percentage points of a multivariate analogue of student's t-distribution. *Biometrika* 42, 256–260

- [14] Dunnet, C. W., and Gent, M. (1977). Significance testing to establish equivalence between treatments, with special reference to data in the form of 2x2 tables. *Biometrics* 33, 593–602
- [15] Dunnet, C.W, Horn, M and Vollandt, R (2001). Sample size determination in step – down and step – up multiple tests for comparing treatments with a control. *Journal of Statistical Planning and Inference* 97, 367 – 384
- [16] Elashoff, J.(2000), nQuery Advisor Release 4.0, Statistical Solutions, Cork, Ireland, Software for MS-DOS systems.
- [17] Freedman, L.S.(1982). Tables of the number of patients required in clinical trials using the Logrank test. *Statist. In Medicine* 1, 121-129.
- [18] Freiman, J.A., Chalmers, T.C., Smith, Jr., H., and Kuebler, R.R.(1986), “The Importance of Beta, the Type II Error, and Sample Size in the Design and Interpretation of the Randomized Controlled Trial: Survey of 71 “Negative” Trials,” in *Medical Uses of Statistics*, eds. J.C. Bailar III and F. Mosteller chap. 14, pp. 289-304, NEJM Boks, Waltham, Mass.
- [19] Freund, J.E., Formerly published under the little Mathematical Statistics, 5th edition, Prentice-Hall, Inc. p 395-396.
- [20] Gail, M., and Gart, J. J. (1973). The determination of sample sizes for use with exact conditional test in 2x2 comparative trials. *Biometrics* 29, 441–448
- [21] George, S.L., and Desu, M.M.(1974). Planning the size and duration of a clinical trial studying the time to some critical event. *J. Chron. Dis.* 27, 15-29
- [22] Gross, A.J., Hunt, H.H., Cantor, A.B., and Clark, B.C.(1987). Sample size determination in clinical trials with emphasis on exponentially distributed responses. *Biometrics* 43, 875-883.
- [23] Grundy, P.M., HEALY, M.J.R. AND REES, D.H. (1956) Economic Choice of the Amount of Experimentation. *J.R. Statis. Soc. A* 18, 32-48.
- [24] Gu, M. and Lai, TL (1999), Determination of power and sample size in the design of cilinical trials with failure – time endpoints and interim analysis. *Controlled Clinical Trials* 20, 423 – 38
- [25] Guenther, W. C. (1977). *Sampling Inspection in Statistical Quality Control*, Macmillan Publishing Co., New York, NY.

- [26] Guenther, W.C.(1979). The use of noncentral F approximations for calculating power and sample size. *Amer. Statistician* 33, 209-210.
- [27] Guttman, I., Wilks, S.S., and Hunter, J.S.(1982) *Introductory Engineering Statistics* (Third Edition), John Wiley and Sons, New York, NY.
- [28] Halperin, M., Rogot, E., Gurian, J., and Ederer, F. (1968). Sample sizes for medical trials with special reference to long-term therapy. *J. Chron. Dis.* 21, 13-24
- [29] Haseman, J. K. (1978). Exact sample sizes for use with the Fisher-Irwin test of 2x2 tables. *Biometrics* 34, 106-109
- [30] Harter, H.L.(1960). Tables of range and studentized range. *Ann. Math. Statist.* 31, 1122-1147.
- [31] Hintze, J.(2000), *PASS 2000*, Number Cruncher Statistical Systems, Kaysville, UT, Software for MS-DOS systems.
- [32] Hochberg, Y., and Lachenbruch, P. A. (1976). Two-stage multiple comparison procedures based on the studentized range. *Commun. In Statist. A5*, 1447-1453
- [33] Hsieh, F. Y. (1989). Sample size tables in Logistic Regression *Statistics in medicine*, 8, 795-802
- [34] Kraemer, H.C. and Theimann, S.(1987), *How Many Subjects? Statistical Power Analysis in Research*, Sage Publications, Newbury Park, CA.
- [35] Koopmans, L., and Qualls, C.(1971). Fixed length confidence intervals for parameters of the normal distribution based on two-stage sampling procedures. *Rocky Mountain J. Math.* 1, 587-602.
- [36] Lachin, J. M. (1981). Introduction to sample size determination and power analysis for clinical trials. *Controlled Clin. Trials* 2, 93-113
- [37] Lakatos, E. (1986). Sample sizes for clinical trials with time dependent rates of losses and noncompliance. *Controlled Clin. Trials* 7, 189-199
- [38] Lee, M. L., Krantowitz, J. L., and Mullis, C. E. (1984). Sample size, cross over trials and the problem of treatment equivalence. *Proceedings of Biopharmaceutical Section. Amer. Statist. Assoc.* 78-80
- [39] Lenth, R.V. (2001), *Some Practical Guidelines for Effective Sample -Size Determination*, Department of Statistics University of Iowa, 1-11.

- [40] Lenth, R.V.(2000), "Java applets for power and Sample size," <http://www.stat.uiowa.edu/~rlenth/Power/>.
- [41] Lieberman, G. J. and, D. B. (1961). Tables of the Hypergeometric Distribution. Stanford University Pres, Stanford, CA.
- [42] Lindley, D.V. (1997) The Choice of Sample Size. The Statistician 46, 129-138.
- [43] Mace, A. E. , (1964). Sample Size Determination. Reinhold Publishing Co., New York.(Reprinted in 1973, by R. E. Krieger Publishing Co. New York)
- [44] Makuch, R.W., and Simon, R.M.(1982). Sample Size Determination. Reinhold Publishing Co., New York. (Reprinted in 1973, by R.E. Krieger Pulishing Co. New York.)
- [45] Morgan, T.M.,(1985). Planning the duration of accural and follow-up for clinical trials. J. Chron. Dis. 12, 1009-1018.
- [46] Moshman, J. (1958). A method for selecting the size of initial sample size in the Stein's two-sample procedure. Ann. Math. Statist. 29, 1271-1275
- [47] Noether, G.E. (1987). Sample size determination for some common nonparametrics tests. J. Amer. Statist. Assoc. 82, 645-647.
- [48] O'Brien, R.G.(1998), UnifyPow. sas Version 98.08.25, Departmant of Biostatistics and Epidemiology, Cleveland Clinic Foundation, Cleveland, OH, Available for download from <http://www.bio.ri.ccf.org/power.html>.
- [49] Odeh, R.E. and Fox, M.(1991), Sample Size Choice: Charts for Experiments with Linear Models, Marcel Dekker, New York, second edn.
- [50] Palta, M., and Amini, S.B.(1985). Consideration of covariates and stratification in sample size determination for survival time studies. J. Chron. Dis. 38, 801-809.
- [51] Pezeshk, H. and Gittins J. Sample Size Determination in Clinical Trials, 1-10.
- [52] Raiffa, H. and SCHLAIFER, R. (1961) Applied Statistical Decision theory. Division of Research Harvard Busines.
- [53] Rodary, C., Comnougue, C., and Tournade, M.-F. (1989). How to establish equivalence between treatments: a one-sided clinical trial in pediatric oncology. Statist. In Medicine 8, 593-598

- [54] Rubinstein, L.V., Gail, M.H., and Santner, T.J.(1981). Planning the duration of a comparative clinical trial with loss to follow-up and a period of continued observation. *J. Chron. Dis.* 34, 469-479.
- [55] Sadooghi-Alvandi, M. (1986). The choice of subsample size in two-stage sampling. *J. Amer. Statist. Assoc.* 81, 555-558.
- [56] Sahai H and Khurshid A (1996). Formulae and tables for the determination of sample size and power in clinical trials for testing difference in proportions for the two – sample design: a review. *Statistics in Medicine* 15, 1 – 21.
- [57] Schoenfeld, D.A.(1981). The asymptotic properties of nonparametric tests for comparing survival distributions. *Biometrika* 65, 316-319.
- [58] Schork, M. A., and Williams, G. W. (1980). Number of observations required for the comparison of two correlated proportions. *Commun. in Statist*, B9, 349–357
- [59] Seelbinder, B. M. (1953). One Stein’s two–stage procedure. *Ann. Math. Statist.* 24, 640–649
- [60] Stein, C. (1945). A two–sample test for a linear hypothesis whose power is independent of variance sample size. *Commun. In Statist.* B16, 957–968
- [61] Thall, P.F., Simon, R., Ellenburg, S.E., and Shraeger, R. (1988). Optimal two-stage designs for clinical trials with binary response. *Statist. In Medicine.* 7, 571-579.
- [62] Thall, P.F., Simon, R., and Ellenburg, S.E. (1988). Two-stage selection and testing designs for comparative clinical trials. *Biometrika* 75, 303-310.
- [63] Thomas, L.(1998), “Statistical power analysis software,” <http://www.forestry.ubc.ca/conservation/power/>.
- [64] Thornley, B. And Adams, C.(1998), “Content and quality of 2000 controlled trials in schizophrenia over 50 years,” *British Medical journal*, 317, 1181-1184
- [65] Wheeler, R.E.(1974), “Portable Power,” *Technometrics*, 16, 193-201.
- [66] Wu, M., Fisher, M., and Demets, D. (1980). Sample sizes for long–term medical trials with time–dependent dropout and event rates. *J. Chron. Dis.* 1, 109–121

ÖZGEÇMİŞ

1980 yılında Almanya'nın Solingen şehrinde doğdu. Lisans eğitimini Mimar Sinan Üniversitesi İstatistik Bölümü'nde 2001 yılında tamamladı. 2002 yılında Ondokuz Mayıs Üniversitesi Sinop Fen-Edebiyat Fakültesi İstatistik Bölümü'nde Araştırma Görevlisi olarak çalışmaya başladı. Aynı yıl, çalıştığı üniversitede Fen Bilimleri Enstitüsü'nde İstatistik Anabilim Dalı'nda yüksek lisans eğitimine başladı. Halen OMÜ Sinop Fen-Edebiyat Fakültesi İstatistik Bölümü'nde Araştırma Görevlisi olarak görev yapmaktadır.

