

Hazırlayan

Göksel ARMAĞAN

ÖRNEKLEME

1.GİRİŞ

Araştırmacının üzerinde durduğu konuyla ilgili popülasyonu tümüyle dikkate alması zaman, emek ve finansman gibi nedenlerden dolayı çoğu zaman mümkün değildir. Bu nedenle araştırmacı popülasyonu temsil eden örnekler üzerinde çalışarak test ve tahminler yapar. Test ve tahminlerin doğru olması için popülasyonla örnek arasındaki ilişkiyi çok iyi kurmak gerekir. Popülasyonla örnek arasındaki ilgi ve örneğin popülasyonu temsil etmesinde dikkat edilecek prensipler Örnekleme Teorisi'nin konusudur.

Belirli bir ana kütle (evrene) veya veri yığınınına uygulanan istatistik teknikler bu verilerin irdelenmesini mümkün kılar. Ancak bütünüyle incelenmesi güç veri yığınlarından ayrılacak belli örneklere aynı tekniklerin uygulanmasıyla ,bu örneklerin seçtikleri veri yığınlarına genellenebilecek sonuçların elde edilmesi mümkündür.

Örnekleme olarak adlandırılan bu durumun en basit tanımı, ana kütle (evren) nin bir kısmı yoluyla ana kütle için genelleştirme yapma işlemidir. Örnekleme ile popülasyon hakkında bilgi edinmek, bu popülasyonu karakterize eden parametreleri tahmin etmek demektir. Örnekleme zaman, finans ve benzeri açılardan daha kullanışlı olduğundan çok geniş uygulama alanına sahiptir. Başta kalite kontrol konusu olmak üzere tüm alanlarda başarılı bir şekilde uygulanmaktadır.

2. ÖRNEKLEMENİN TANIMLANMASI VE KULLANILMASININ NEDENLERİ

Bir ana kütle (evren) den seçilmiş ve daha az sayıda birimden oluşan örneği gözlemleyerek elde edilen istatistiksel sonuçların tüm ana kütle (evren) için genelleme yapılmasına **örnekleme** denir.

Örnekleme kullanmanın nedenleri:

- Belirli testlerin yıkıcı doğası
- Ana kütle (Evren) deki tüm elemanların kontrol edilmesinin fiziksel imkansızlığı
- Bir ana kütle (evren) deki tüm elemanlarla çalışmanın maliyetinin yüksekliği
- Örnek sonuçlarının yeterliliği
- Tüm ana kütle (evren) yi incelemedeki zaman kaybı
- Hızlı sonuç alma
- Daha büyük çalışma alanı
- Daha büyük doğruluk

3. ÖRNEKLEME İLE İLGİLİ BAZI TEMEL KAVRAMLAR :

Ana kütle (Evren) : Gözlemlenecek olan tüm birimlerin oluşturduğu kütleye ana kütle (evren) denir. Ana kütle (Evren) ye aynı zamanda popülasyon ya da yığın denilmektedir.

Örnekleme: Ana kütle (Evren) den gözlemlenmek üzere seçilen bir kısım elemanın oluşturduğu küçük grup olarak adlandırılır.

Örnekleme Hacmi: Örneklemenin oluşturduğu elemanlar örnekleme hacmi olarak ifade edilir. Ve “n” ile sembolize edilir.

Çerçeve: Gözlemlenecek ana kütlede bir kısım elemanları rastgele seçebilmek için ana kütle (evren) deki tüm elemanların bir liste halinde düzenlenmesi gerekir. Ana kütle (Evren) deki tüm elemanları içeren ve bunların birbirinden bağımsız olarak belirlenmesini olanaklı kılan liste veya listeye benzer herhangi bir araç örneklemin çerçevesini oluşturur. Çerçevenin düzenlenmesinde çerçeve maliyeti, sistematik hata ve elemanların belirli sayıya indirilmesi gibi kısıtlamaların dikkate alınması çerçevenin hata payı ve maliyetinin düşük olmasını sağlar.

Ana kütle (Evren) yi oluşturan elemanlar, tamsayıda istatistik elemanı, örneklemede örnekleme elemanı olarak adlandırılır. Örnekleme elemanı belirlenerek tanınması kolay, örnekleme tipine uygun ve düşük maliyet gerektirmesine özen gösterilmelidir. Gözlem elemanı ise örnekleme elemanının bir parçasını oluşturur.

4. ÖRNEKLEME TEKNİKLERİ

Örnekleme türleri ile ilgili olarak çeşitli sınıflandırmalar yapılabilir. Örnekleme türlerinin olasılıklı ve olasılıksız olmak üzere iki şekilde ele alınarak sınıflandırmanın yaygın olarak kullanıldığı söylenebilir. Bu kısımda ele alınacak olan olasılıklı ve olasılıksız örnekleme yöntemleri aşağıda belirtilmiştir.

I. Olasılıklı Örnekleme Yöntemleri:	II. Olasılıklı Olmayan Örnekleme
<ul style="list-style-type: none">• <i>Basit rastgele örnekleme</i>• <i>Sistematik örnekleme</i>• <i>Tabakalara göre rastgele örnekleme</i>• <i>Kümelere göre örnekleme</i>	<ul style="list-style-type: none">• <i>Kota örnekleme</i>• <i>Yargılı örnekleme</i>• <i>Kolaylık örnekleme</i>

4.1. Olasılıklı Örnekleme Yöntemleri

N elemandan meydana gelen bir ana kütle (evren) den “n” elemanlı örnekleme yapıldığında olasılık esaslarına göre seçim yapıldığında seçim yöntemleri şöyle sıralanabilir.

Rastgele Rakamlar Çizelgesi İle Seçim:

Örnekleme yapılan ana kütle (evren) deki eleman sayısı çok fazladır. Bu nedenle her bir elemanın bir belgesinin olması ve seçimin bu esasa göre yapılması zordur. Bu durumlarda uygulanması kolay olan bir yöntem kullanılır. Ana kütle (Evren) deki her elemana sıra numarası verilir fakat örneğe girecek elemanların sayısı çok olduğu durumlarda rakamların tek tek düzenlenmesi zaman alır. Bu dezavantajı önlemek için İngiliz istatistikçi Tippett tarafından “rastgele rakamlar çizelgesi” düzenlenmiştir.

Söz konusu bu çizelgenin hazırlanmasında , 0000, 0001,9998, 9999 rakamlarından tesadüfî olarak seçilenler sıraya göre işlenerek bir listenin oluşturulması prensibi temel alınmıştır.

İadeli ve İadesiz Seçim:

Ana kütle (Evren) den örnekleme yapılırken ana kütle (evren) elemanın aynı örneğe birden fazla girebilmesine iadeli seçim denir. İadeli seçimde her elemanın her çekilişte örneğe girme olasılığı $1/N$ 'dir. Çekilişler “n” defa tekrarlandığından bir elemanın örneğe girme olasılığı n/N 'dir.

İadesiz seçimde bir elemanın herhangi bir çekilişte örneğe girme olasılığı $1/N$ olduğundan hareketle ana kütle (evren) deki herhangi bir istatistik elemanın “n” elemanlık bir örneğe girme olasılığı n/N 'dir.

4.1.1. Basit Rastgele Örnekleme:

Basit rastgele örnekleme, “N” büyüklüğündeki bir ana kütle (evren) den bağımsız bir şekilde seçilecek “n” büyüklüğündeki örneklerin eşit şansa sahip olarak belirlenmesini sağlayan bir yöntemdir. Tanımdan da anlaşılacağı gibi ana kütle (evren) deki her bir parça veya kişi aynı şansa sahiptir. Rastgele bir örneğin seçilmesinde en iyi yöntem ana kütle (evren) deki her bir elemana bir tanımlama numarası vermek ve rastgele sayılar çizelgesini kullanmasıdır. Bu seçim rastgele süreç yardımıyla yapılır. Sayının her bir rakamı için 0, 1, 2,....., 9'un olasılığı aynıdır. Böylece ön yargı, seçimin yapısından dolayı tamamen ortadan kaldırılmış olur.

4.1.2. Tabakalara Göre Rastgele Örnekleme

Bir ana kütle (evren) nin üst üste çakışmayan gruplara bölünmesi tabakalandırma olarak adlandırılır. Bir ana kütle (evren) ilk olarak tabakalandırılır ve bunu takiben her bir tabakadan bir rastgele örnek seçilirse, izlenen yol tabakalandırılmış rastgele örnekleme olarak adlandırılır. Örneklerin tüm ana kütle (evren) den seçilmesinin yerine her bir tabakadan seçilmesinin bir nedeni , her bir tabakadan belirlenen örneklemelerin yeterli büyüklükte olmasından emin olunmasıdır.

Tabakalı Rastgele Örneklemenin Kullanılmasının Sebebi:

- Her bir tabaka için yeterli örneklemenin yapılabilmesi,
- Diğer örnekleme tiplerine nazaran ana kütle (evren) karakteristikleri konusunda daha kesin tahminlemenin sağlanabilmesi.

Örnekleme sonlu bir anakütleden iadesiz olarak yapılacaksa örnek hacmi

$$n = \frac{Nz^2\sigma^2}{d^2(N-1) + z^2\sigma^2}$$

formülü ile bulunur.

Burada:

N=Anakütle büyüklüğü

σ^2 =Anakütle varyansı

z=Güven düzeyi

d=Hata değeri

Güven düzeyi, örneğin %99 (çift taraflıdır) için 2.58; %95 (çift taraflıdır) için, 1.96; %90 (çift taraflıdır) için, 1.645'dir.

Örneğin z tablosundan %95 için, 0,95 değeri ikiye bölünerek 0.475 değeri elde edilir ve 0.475 değeri tablodaki üst sol standart sapma değerleri kesştirilerek 1.96 değeri bulunur.

Hata değeri tahmin etmek istenilen gerek aritmetik ortalamaya ne kadar yaklaşmak istediğimizi ifade eder. Bir başka ifadeyle, gerçek anakütle parametresi için hata değeridir.

4.1.3. Sistematik Örnekleme

Bireyleri 1'den N'e kadar sıralanmış anakitlede, sistematik olarak her k'ncı bireyi seçmek suretiyle yapılır. Bir süpermarkete alışverişe gelenlerden örnekleme yapmak istenildiğinde, örnek hacmi 50 müşteri ve süpermarketin günlük ortalama müşteri sayısı 500 olsun. Bu durumda her $500/50=10$ kişiden birini örneğimize seçebiliriz. 1 ile 10 arasında tesadüfi bir sayı çektiğimizi ve bu sayının 4 olduğunu kabul edersek, örneğe giren ilk birey, o markete giren 4. Müşteridir. Takip eden bireyler $4+10=14$. Müşteri, $14+10=24$. Müşteri, $24+10=34$. Müşteri şeklinde devam edecektir.

4.1.4. Kümelere Göre Örnekleme

Araştırmalarda büyük bir kent, coğrafi bölge ya da ülke üzerinde çalışılması gereken durumlarda ana kütleli oluşturan listenin elde edilmesi en önemli sorundur. İkinci sorun ise, ana kitle listesi elde edilse bile, çok geniş bir alana yayılarak çalışmanın getirdiği zorluktur. Böyle durumlarda araştırmacılar kümeli (cluster) örneklemeden yararlanabilir. Kümeli örneklemede, anakitle küme adı verilen N adet gruba bölünür. Örneğin, bir kent semtlere veya gelir düzeyine göre bölgelere ayrılır. Daha sonra bu kümelerden tesadüfi bir örnek seçilir ve örneğe giren kümelerdeki tüm bireylerle görüşülür. Kümelerden elde edilen veriler birleştirildikten sonra, anakitle ile ilgili tahminlemede bulunulur.

4.2. Olasılıklı Olmayan Örnekleme Yöntemleri

Örnek büyüklüğünün belirlenmesinde örneğe girecek elemanların oluşturulmasından ve örnek sonuçlarından hareketle ana kütle (evren) için tahminlemede araştırmanın subjektif görüş ve eğilimini temel alan iradi metotla karar verilmesi durumu olasılıklı olmayan örnekleme olarak tanımlanır. Kullanılan yöntemler şunlardır.

4.2.1. Kota Örnekleme

Bu yöntemde araştırmacının subjektif görüş ve eğilimlerinin sonuçları etkilememesi için değişik kişilere belli bir özelliğin analizlerinin verilmesidir. Bu özellik bir kota oluşturur. Bu yöntemde ana kütle (evren) nin kotalara göre dağılışı biçiminin önceden bilinmediği kabul edilir.

4.2.2. Yargı Örnekleme

Yargı ve karar tamamen araştırmacının kendisine aittir. Bu araştırma çok deneyimli araştırmacılara yaptırılır. Örneğin; Ege bölgesi süt ve süt ürünleri hakkında bir anket yapacaksa Ege bölgesini iyi bilen birisi yapmalıdır.

4.2.3. Kolaylık Örnekleme

Anakitleyi tam içermeyen listelerden yararlanılarak seçimin yapıldığı örnekleme metodudur. Örneğin telefon rehberinden örnek fert seçmek.

Örnekleme Hacminin Belirlenmesine İlişkin Örnekler

Örnek 1:

Bir ilçede toplam 2500 tarım işletmesi bulunmaktadır. Bu işletmelerin arazi büyüklüğünün varyansı 9 dekadardır. 0.5 ve 1.0 dekar hata, %95 ve %99 güvenle örnek hacmini hesaplayınız.

$$n = \frac{Nz^2\sigma^2}{d^2(N-1) + z^2\sigma^2} \text{ kullanılır. Buna göre;}$$

a) %95 için; $z=1.96$, $d=0.5$

$$n = \frac{Nz^2\sigma^2}{d^2(N-1) + z^2\sigma^2} = \frac{2500(1.96)^2(9)}{0.25(2499) + (1.96)^2(9)} = \frac{86436}{659.32} = 131.09 \cong 132$$

b) %95 için; $z=1.96$, $d=1.0$

$$n = \frac{Nz^2\sigma^2}{d^2(N-1) + z^2\sigma^2} = \frac{2500(1.96)^2(9)}{1(2499) + (1.96)^2(9)} = \frac{86436}{2533.57} = 34.12 \cong 34$$

c) %99 için; $z=2.58$, $d=0.5$

$$n = \frac{Nz^2\sigma^2}{d^2(N-1) + z^2\sigma^2} = \frac{2500(2.58)^2(9)}{0.25(2499) + (2.58)^2(9)} = \frac{149769}{684.66} \cong 219$$

d) %99 için; $z=2.58$, $d=1.0$

$$\frac{2500(2.58)^2(9)}{1(2499) + (2.58)^2(9)} = \frac{149769}{2558.9} \cong 59$$

Görüldüğü gibi güven düzeyi yükseldikçe, örnek hacmi artmakta, hata değeri arttıkça, örnek hacmi azalmaktadır.

Örnek 2:

30000 hanehalkı bulunan bir kent merkezinde yeni bir ürün için potansiyel talep belirlemek üzere, hanehalkı anketleri yapılacaktır. Bu amaçla maksimum anket sayısının ne olması gerektiğini %5 ve %10 güven aralıkları için 0.05 ve 0.10 hata paylarına göre belirleyiniz.

N büyüklüğündeki sonlu bir anakütle için, belli bir özelliği taşıyanların bilinen veya tahmin edilen oranına (p) göre örnekleme yapılmak isteniyorsa,

$$n = \frac{Np(1-p)}{(N-1)\sigma_{p_x}^2 + p(1-p)} \text{ olur. Burada } p, \text{ üzerinde çalıştığımız özelliğin}$$

anakütledeki oranı veya rastlanma olasılığıdır. P, daha önce araştırma veya deneyimlerden elde edilebileceği gibi, sezgisel olarak da tahmin edilebilir. Maksimum örnek hacmi için $p=0.5$ alınmalıdır. p'nin 0.5'ten daha az ve daha yüksek değerleri örnek hacmini düşürür. O nedenle, p'nin bilinmediği durumlarda maksimum örnek hacmiyle çalışmak olası hatayı azaltacağından $p=0.5$ alınması önerilir. $\sigma_{p_x}^2$ ise, oranın varyansıdır. Gerçek oran ne olursa olsun, bunun istenen herhangi bir olasılık düzeyinde güven aralığının, örnek oranının iki tarafında belli bir r oranından daha fazla uzanmaması istenebilir. Bu durumda $\sigma_{p_x}^2$ parametresini:

$Z_{\alpha/2}\sigma_p = r$ formülü ile elde edebiliriz. Örneğin:

- Anakütle oranına ait %95 güven aralığının, örnek oranının 0.05 iki tarafında uzanmasını istiyorsak:

$$Z_{\alpha/2}\sigma_p = r \quad 1.96\sigma_p = 0.05 \quad \sigma_p = \frac{0.05}{1.96} = 0.0255$$

- Anakütle oranına ait %95 güven aralığının, örnek oranının 0.10 iki tarafında uzanmasını istiyorsak:

$$1.96\sigma_p = 0.10 \quad \sigma_p = \frac{0.10}{1.96} = 0.05102$$

- Anakütle oranına ait %90 güven aralığının, örnek oranının 0.05 iki tarafında uzanmasını istiyorsak:

$$1.645\sigma_p = 0.05 \quad \sigma_p = \frac{0.05}{1.645} = 0.03039$$

- Anakütle oranına ait %90 güven aralığının, örnek oranının 0.10 iki tarafında uzanmasını istiyorsak:

$$1.645\sigma_p = 0.10 \quad \sigma_p = \frac{0.10}{1.645} = 0.06079 \text{ olur.}$$

σ_p büyüdükçe, örnek hacmi azalır. Eğer amaç örnek hacmini azaltmaksa, güven aralığını daraltıp, hata payını artırmak gerekir.

Örnek çözümü için $n = \frac{Np(1-p)}{(N-1)\sigma_p^2 + p(1-p)}$ formülünde;

$$N=30000$$

Maksimum örnek hacmi istendiğinden $p=0.5$, güven aralıkları ve hata paylarına göre sırasıyla;

a) %95 güven aralığı, 0.05 hata payı için:

$$1.96\sigma_p = 0.05 \quad \sigma_p = 0.02551, n = \frac{30000(0.5)(0.5)}{(29999)(0.02551)^2 + (0.5)(0.5)} \cong 379$$

b) %95 güven aralığı, 0.10 hata payı için:

$$1.96\sigma_p = 0.10 \quad \sigma_p = 0.05102, n = \frac{30000(0.5)(0.5)}{(29999)(0.05102)^2 + (0.5)(0.5)} \cong 95$$

c) %90 güven aralığı, 0.05 hata payı için:

$$1.645\sigma_p = 0.05 \quad \sigma_p = 0.03039, n = \frac{30000(0.5)(0.5)}{(29999)(0.03039)^2 + (0.5)(0.5)} \cong 268$$

d) %90 güven aralığı, 0.10 hata payı için:

$$1.645\sigma_p = 0.10 \quad \sigma_p = 0.06079, n = \frac{30000(0.5)(0.5)}{(29999)(0.06079)^2 + (0.5)(0.5)} \cong 68$$

KAYNAKLAR

1. ARIKAN, R. (2013). Araştırma Yöntem ve Teknikleri, 2.Basım, Nobel Yayıncılık, ANKARA.
2. Miran, B. (2013). Temel İstatistik, İZMİR.