

BETİMSEL İSTATİSTİK NOT

İSTATİSTİK DERS NOTU

1. TEMEL KAVRAMLAR

İstatistiğin Tanımı ve Kullanım Alanları

İstatistik Nedir?

Genel anlamıyla istatistik , yaşama ilişkin düşüncelerin sistematize edilmiş biçimidir. İstatistik iş yaşamında da, işletmedeki her türlü faaliyetin kantitatif (nicel) temellerini hazırlar. Bu bakış açısıyla istatistik; “daha etkin kararlar vermek amacıyla verileri toplama, sınıflama, sunma, çözümleme ve yorumlama biçimi” olarak tanımlanabili

Diğer bir tanıma göre istatistik, verilerin toplanması, işlenmesi, analiz ve yorumunda kullanılan metodlar bütünüdür

İstatistiğin Kullanım Alanları

İstatistik günlük hayatımızın hemen her alanında kullanılmaktadır. Aşağıda istatistiğin kullanım alanlarından bazılarına yer verilmiştir.

- ☐ IMKB 'de hisse senetlerinin analizi
- ☐ Siyasi partilerin seçimlere ilişkin kamuoyu yoklamaları
- ☐ Kalite Kontrol
- ☐ Şirketin muhasebe kayıtlarının denetlenmesi
- ☐ Pazarlama araştırmaları
- ☐ Ekonomik göstergelerin takibi
- ☐ Tıbbi araştırmalar
- ☐ Mühendislik araştırmaları
- ☐ Bilimsel çalışmalar
- ☐ Uzay araştırmaları
- ☐ Demografik (Nüfus özellikleri) araştırmalar

Temel Kavramlar

Araştırma

İlgilenilen konuya ilişkin sorunların saptanması, çözüm yollarının planlanması, uygulamaya konulması ve sonuçlarının değerlendirilmesine yönelik yapılan çalışmalardır. Araştırmanın konusu ve amacı belirgin olmalıdır. Zaman, personel ve maliyet dikkate alınarak, araştırmanın sınırı iyi belirlenmelidir. Araştırmada görev alan araştırmacılar ya da uzmanlar yeterli bilgi düzeyine sahip olmalıdır. Araştırmanın sonucunda elde edilen bilgiler doğru biçimde değerlendirilmelidir

Popülasyon (Kitle, Ana Kitle, Ana Kütle, Yığın)

Araştırma kapsamına giren, aynı özellikleri taşıyan birimlerin ya da bireylerin oluşturduğu topluluğa KİTLE denir. Kitlenin büyüklüğü araştırmanın özelliğine göre değişir. Nüfus sayımı için kitle Türkiye’dir. Denizli’deki üniversite öğrencilerinin giderleri için kitle Pamukkale Üniversitesi öğrencileridir

Örneklem ve Örnekleme

Kitle büyüklüğüne bağlı olarak her zaman tüm birimler (bireyler) hakkında bilgi sahibi olmak mümkün değildir. Bundan dolayı geniş kitlelerde araştırmalar; zaman, maliyet, personel, ulaşım, vb. nedenlerden dolayı tüm birimler yerine daha az sayıdaki birimler seçilerek yürütülür

Bir kitleden, belirli yöntemler kullanılarak seçilen aynı özellikleri taşıyan bir kısım bireyin oluşturduğu topluluğa ÖRNEKLEM denir

Örneklem; ana kitleyi nitelik ve nicelik yönünden temsil edebilecek bir kümenin çekilmesi işlemidir

Tamsayım

Bir araştırma kapsamında, kitledeki tüm birimlerine ulaşarak istenen bilginin elde edilmesi işlemidir. Bunun yapılabilmesi için incelenecek kitlenin büyüklüğünün, belirlenen maliyet ve zaman gibi kısıtlara uygun olması gerekir. Bazı durumlarda (nüfus sayımları gibi) kitle büyük olsa bile tam sayım yapılması zorunlu olmaktadır. Gelişen teknoloji ile birlikte bu tür tam sayımlar daha kolay yapılabilir hale gelmiştir

Gözlem (Denek)

Kitle ya da örneklemde yer alan her birime gözlem ya da denek denir. Gözlem (ya da denek) sayısı aşağıdaki biçimde simgeleştirilmektedir.

Kitledeki Gözlem Sayısı : N

Örneklemdeki Gözlem Sayısı: n

Parametre ve İstatistik

Kitle özelliklerinin sayısal değerlerine PARAMETRE denir. Araştırma kitle yerine örneklem üzerinde uygulanıyorsa, parametre değerleri tahmin edilir. Bu durumda, örneklemde elde edilen sayısal değerlere İSTATİSTİK denir. Örnek olarak, sıkça kullanılan bazı parametreler ve istatistikler aşağıda verilmektedir:

Parametre İstatistik

Kitle Ortalaması : μ Örneklem Ortalaması : $\bar{x}$

Kitle Varyansı : σ Örneklem Varyansı : S^2

Değişkenler

Değişkenin kelime anlamı; Değişme özelliği gösteren, çok değişen, değişebilir, kararsız, değişici şeklindedir. Matematiksel tanımı ise; Gözlemde gözleme değişik değerler alabilen objelere, özelliklere ya da durumlara "Değişken" denir.

İstatistikte değişken(variable) terimi, deneklere ait özellikler anlamında kullanılır. Örneğin araştırmaya katılanların cinsiyeti, medeni hali, gelir düzeyi, eğitim düzeyi, coğrafik konumu gibi demografik özellikleri ve boyu, kilosu, vizeden aldığı not, sorulara verdikleri cevaplar gibi özellikler birer değişkeni ifade ederler.

Nicel (Kantitatif) Değişkenler

Birimlerin ölçüm ve tartım sonucu değerleri saptanan sayısal özelliklerini belirten değişkenlerdir. Bu değişkenler değerleri, mekanik ve elektronik araçlara sayısal olarak aralıklı ölçekli yada orantılı ölçekli verileridir. Nicel değişkenlerin verilerine nicel veri adı verilir.Örneğin birimlerin, boy uzunluğu, vücut ağırlığı, kilosu, kan basıncı gibi özellikler nicel değişkenlerdir.

Kesikli (Süreksiz) Değişken

Bu değişkenler miktar yönünden değişiklik yerine tür yönünden değişiklik gösterir. Dolayısıyla bir obje ya da birey bir özelliğe sahiptir ya da değildir. Yani kesin değerler alırlar. Nitel değişkenlerin hemen hepsi süreksiz değişkendir. Cinsiyet, medeni durum, göz rengi gibi. Birinin diğerine göre daha çok veya az olması mümkün değildir.

Sürekli Değişken

İki ayrı ölçüm arası kuramsal olarak sonsuz parçaya bölünebilir. Ölçüm söz konusu olduğu için sürekli değişken değerleri her zaman tam değeri vermezler. Rasyonel sayılar kümesinin elemanları ile belirtilirler.Yaş, uzunluk ve ağırlık gibi

Ölçme tartma yoluyla elde edilen dolayısıyla nokta içermesi mümkün olan verilerdir. Örneğin; balıkların ağırlıkları, tohum ağırlıkları, tohum çapı, bitki boyu, ortamdan bakterilerin tüketmiş oldukları şeker miktarı, ineklerin yıllık süt verimleri gibi.

Nitel (Kalitatif) Değişkenler

Nitel değişkenler; birimlerin kalite, kategorik, yada isimsel olarak belirtilebilen karakteristik özelliklerini, durumlarını ve pozisyonlarını belirten değişkenlerdir. Bu değişkenlerin verileri isimsel ya da sıralı ölçekle elde edilmişlerdir ve iki yada daha fazla kategoriye (alt seçenek, sınıf, grup) ayrılarak sayımla elde edilir. Bu değişkenlerin verilerine nitel veriler adı verilir. Örneğin birimlerin, cinsiyeti, kan grubu, medeni durum, göz rengi, mesleği, yerleşim yeri, tuttuğu futbol takımı (fanatikler için) gibi nitelik bildiren durumları açıklayan değişkenlerdir.

Değişkenlerin Ölçümü

İstatistik analize başlamadan önce ilk yapılacak şey, değişkenlerin nasıl ölçüldüğünün belirlenmesidir. Bu sorunun yanıtı "sfıgmomanometre, Sahli yöntemi, termometre, otoanalizer, ..." gibi cihaz ya da yöntemler ya da "santimetre, yıl, IU/mL" gibi birimler olabilir; ancak istatistik biliminde "ölçüm"den kastedilen, bu değildir. İstatistikte ölçüm denildiği zaman anlaşılan, değişkenin alabileceği değerlerle ilgili kısıtlamalardır. Örneğin bir insanın cinsiyeti "bayan" yada "bay" değerini, yaşı 1, 2, 3 ... değerlerini, boyu 1.75 yada 1.85 değerlerini alabilir.

Değişkenlerin alabileceği değerlerin neler olabileceği, neler olamayacağı, yani nasıl ölçüldüğünü belirlemek, yapılacak istatistik analizin seçimi için çok önemlidir. Ölçüm özelliklerine göre değişkenleri başlıca üç gruba ayırabiliriz: nominal, ordinal ve sayısal değişkenler.

Kullanılan Ölçek Türleri

Nominal Ölçek

Nominal ölçekte veriler sadece niteliklerine göre isimlendirilip gruplandırılabilir. Bu ölçekle ölçülen veriler ile matematiksel işlem yapılamaz. Cinsiyet, medeni durum, din , parça numarası, çalışma durumu, meslek vb. kalitatif veriler nominal ölçekle ölçülürler. Bu verilerin frekansları ve modları belirlenebilir

Nominal bir değişkende ölçüm düzeyleri arasında bir sıralama ya da uzaklık-yakınlık gibi belirli bir mesafe yoktur.

Ordinal Ölçek

Bu ölçekte veriler nominal özellik taşımakla birlikte aynı zamanda sıraya konulabilir

Ordinal bir değişkende ölçüm düzeyleri arasında bir sıralama vardır, ama düzeyler arasındaki mesafeler belirli değildir.

Bu ölçeğin amacı, bir konu hakkındaki düşünceleri belirli bir öncelik sırasına koymaya hizmet etmektir. Bu ölçek, belli bir özelliğe göre nesnelerin yada yargıların konumlarını, aralarındaki uzaklığı belirtmeksizin dile getiren bir ölçek türüdür

Örneğin restaurant, otel, hastane gibi hizmet işletmeleri yiyecek kalitesi, servis, tesisler, konukseverlik gibi özellikler açısından “mükemmel”, “iyi”, “orta”, “kötü”, “çok kötü” olarak sınıflandırılabilirler. Bu sıralamada “mükemmel” için 1, “iyi” için 2, “orta” için 3, “kötü” için 4, “çok kötü” için 5 değeri kullanıldığında; “mükemmel” olarak değerlendirilen bir A otelinin “iyi” olarak değerlendirilen bir B otelinden iki kat iyi olduğu söylenemez. Bu ölçekle ölçülen verilerin frekansı, modu, medyanı, yüzdelik değeri ve korelasyonu hesaplanabilir

Aralık Ölçeği

Bu ölçekte veriler ordinal özellik taşımakla birlikte, aynı zamanda bu veriler için ölçekte tanımlanan birimlerin birbirinden eşit uzaklıkta bulunduğu bir ölçü birimi ve rastgele bir başlangıç noktası tanımlanabilir

Örneğin su buz haline 0, kaynama haline 100 değeri verilerek aradaki ısı farkı 100 eşit parçaya bölünüp her aralığa sırası ile bir derece verilir. Ancak burada dikkat edilmesi gereken nokta şudur: Bu ölçekle ölçülmüş iki nesneden biri 10, öteki 20 puan almış ise, biri ötekinin yarısı yada iki katı değildir

Bu ölçekle ölçülen verilerin aritmetik ortalama, standart sapma, yatıklık ve basıklık ölçüleri hesaplanabilir

Oran Ölçeği

Bu ölçek, objektif bir başlangıç noktasını temel alan ve üzerinde ölçülecek nokta veya objeleri bir birinin katı olarak ifade eden bir ölçektir

Bu ölçek, aralıklı ölçeğin özelliklerini taşımasının yanı sıra bu verilerden iki ölçümlemenin oranlanmasıyla anlamlı kıyaslamalar elde edilir. Uzunluk, parasal değerler, zaman, güç değerleri oran ölçeğiyle ölçülürler

Ölçmede Dikkat Edilmesi Gereken Hususlar, Ölçek Güvenilirliği ve Ölçek Geçerliliği

Toplumbilimde bağımlı ve bağımsız değişkenleri ölçmek her zaman kolay olmadığı için, kullanılan ölçeklerin istenen bilgiyi toplamaya elverişli olup olmadığı büyük önem taşır. Bu nedenle, ölçekleri hem seçerken hem de kullanırken iki önemli noktaya dikkat edilmelidir

- ☐ Kullanılan ölçek, istenen bilgiyi tam, doğru ve objektif olarak ölçebilecek mi?
- ☐ Kullanılan ölçek, araştırmanın hipotezlerini test etmeye hizmet edecek mi? Ve uygulanması düşünülen test tekniklerine elverişli mi?

Bir ölçeğin güvenilirliği, onun sağlamlığını ve tutarlılığını ifade eder. Bir göstergenin ölçüm aracının kendi yüzünden çeşitlilik göstermemesi, kullanılan ölçeğin güvenilir olduğunu gösterir

Güvenilirlik zaman, temsil ve eşdeğerlik açısından üç ayrı özellik taşır. Zaman özelliği, aynı ölçeğin farklı zamanlarda uygulanması sonucunda aynı sonuçları vermesidir. Temsil, aynı ölçeğin farklı gruplara uygulanması halinde aynı sonuçları vermesi demektir. Eşdeğerlilik, ölçümler arasında tutarlılığın sağlanıp sağlanmadığını ifade eder

Geçerlilik, kullanılan ölçeğin, ölçülmek isteneni tam olarak ölçüp ölçmemesine bağlıdır

İstatistik Türleri

İstatistik teknik ve yöntemlerini kullanma amaçlarına göre iki gruba ayırabiliriz.

- ☐ Tanımlayıcı (Betimleyici) İstatistik
- ☐ Öngörüleyci (Tahminleyici) İstatistik

Tanımlayıcı (Betimleyici) İstatistik

Tanımlayıcı istatistik sayısal verileri sınıflama ve özetlemede kullanılan yordamlardır. Verileri tablo, grafik veya sayısal olarak anlamlı bir biçimde özetler. Bazı veriler frekans dağılımı olarak düzenlenebilir

verilerden ortalama değer ve bazı özel orta değerler hesaplanabilir. Örneğin, medyan bir grup sayısal veriyi ikiye bölen (%50 - %50) orta noktadaki değerdir

Öngörüleyci (Tahminleyici) İstatistik

Öngörüleyci istatistik, gözlemlenmiş durumlardan elde edilen verilerle, gözlemlenmemiş durumlar için sonuç çıkarır ve popülasyon hakkında öngörüleme yapar. Popülasyon bireylerden oluşabilir; Celal Bayar Üniversitesi'nde okuyan öğrenciler, istatistik 1 dersini alan öğrenciler, Sağmalcılar ceza evindeki tutuklular gibi. Popülasyon objelerden de oluşabilir: Ford fabrikasında geçen hafta üretilen Ford Ka arabaları gibi. Popülasyon bir grup ölçümden oluşabilir: Efes Pilsen Basketbol takımındaki oyuncuların boyları vb

Veri Kaynakları

İstatistiksel çalışmalarda, ihtiyaç duyulan veriler, iç ve dış kaynaklardan elde edilirler.

İç Kaynaklar

Genelde örgütlerde, özelde de işletmelerde gerçekleştirilen istatistiksel araştırmalar için gerekli veriler çoğunlukla örgütün yada işletmenin kayıtlarından elde edilir. İşletmenin muhasebe, üretim, stok, bordro, satış ve müşteri hesapları kayıtlarından alınan verilerine iç veriler denir. İç verilerin toplanması genelde sorun yaratmaz

Dış Kaynaklar

İç veriler genellikle önemli olsalar da birçok durumda dış kaynaklardan bazı bilgilerle desteklenmelidirler. Devletten, belediyelerden, bankalardan, ticaret ve sanayi odalarından ve diğer özel kurumlardan elde edilen verilere dış veriler denir. Dış kaynaklardan veri toplanırken bazı sorunlarla karşılaşılır. İki ayrı dış kaynaktan elde edilen veriler birbirine uymayabilir yada istenilen maliyet ve süre içinde uygun dış veri bulunmayabilir

Birincil Veriler

Araştırmaya konu olan ana kitleden doğrudan elde edilen bilgiler yada verileri toplayan kuruluşlardan elde edilen veriler birincil veri olarak tanımlanırlar. Bazı bilgiler ikincil kaynaklardan ve başka örgütlerden bulunamadığından, bunları doğrudan birincil kaynaklardan toplama zorunluluğu vardır. “Tüketicilerin otomobil satın alma kararlarında aile reisi dışındaki aile bireylerinin bu karar üzerinde etkili olup olmadıkları” araştırılacaksa birincil kaynaklara yönelmek gerekir. Yani, bu bilgiler ilgili bireylerden elde edilecekti. Genellikle birincil kaynaktan sağlanan veriler tercih edilmektedir. İkincil kaynaktan yayın hataları olabileceği gibi birincil kaynaktan veriler düzgün ve tam tanımlamalarla verilir. DPT, DİE bültenleri ve yayınları TÜFE ve TEFE’ yi kendileri hesaplayarak yayınladıkları için birincil verileri içerirler.

İkincil Veriler

Verileri toplayan örgütten başka bir örgüt tarafından yayınlanmış veriler ise ikincil veri olarak tanımlanır. Eğer TÜFE, gazetelerin finans sayfalarından elde ediliyorsa ikincil veri sayılır (

Veri Toplama Yöntemleri

Birincil ve ikincil veri kaynaklarının araştırılması ve verilerin toplanması için kullanılan bazı yöntemler mevcuttur. Bu yöntemler; literatür araştırması, gözlem, deney ve anket yöntemleridir.

Literatür Araştırması

Araştırma konusu ile ilgili benzer bilimsel araştırmalar ve raporlar, indeksler vb. kaynaklar bilimsel literatürde, ilgili kuruluşların sitelerinde detaylı olarak araştırılmalıdır. Bu şekilde bulunabilecek kaynaklardan veriler hazır olarak elde edilebileceği gibi, veri toplama araçlarının tasarımı için faydalı bilgiler de elde edilebilir.

Gözlem

Gözlem; kendiliğinden oluşan ya da bilinçli olarak hazırlanan olayları, belirdikleri sırada sistemli ve amaçlı bir biçimde inceleyerek bilgi toplama yöntemidir. Biz mezhebe mensup insanların dini törenlerini incelemek için bilgi toplayan bir araştırmacı, o mezhebe mensup insanların dini ayinlerine katılarak ya da bu ayinleri izleyerek bilgi toplayabilir

Gözlem Yönteminin Avantajları:

- ☐ Anket yöntemine göre daha objektiftir,
- ☐ Hızlı bilgi toplamaya olanak sağlar,
- ☐ Uygulanması kolaydır.

Gözlem Yönteminin Dezavantajları:

- ☐ Her tür bilginin toplanmasına uygun değildir,
- ☐ Mekanik ve elektronik araçların kullanılmaması halinde, olayın ayrıntılarını yakalamak veya hatırlamak zordur,
- ☐ Gözlem altında olduklarını anlayanlar, olağan davranışlarını değiştirebilirler,
- ☐ Bazı durumlarda maliyetleri yüksektir.

Deney

Deney, bir hipotezin sınanması amacı ile koşulları deneyi yapan tarafından hazırlanan ve bağımsız değişkenin bağımlı değişken üzerindeki etkinin şiddetini ya da yönünü ortaya koymayı amaçlayan bir gözlem türüdür

Anket

Anket, birincil kaynaklardan bilgi toplamak için hazırlanan sistematik bir soru formudur. Amacı, araştırmanın problemini çözecek ve ele alınan hipotezleri test edecek bilgileri sistematik bir biçimde toplamak ve saklamaktır

Anketlerin Uygulanma Türleri

Yüz Yüze Görüşme

Anketörle deneklerin aynı fiziki ortamda bulunduğu, anket sorularının doğrudan sözlü olarak yöneltilmesi ve yine sözlü olarak alınan cevapların anket formuna not edilmesi şeklindeki bir veri toplama yöntemidir. Yöntemin avantajlar

- ☐ Bu yöntem hem soruları açıklığa kavuşturma hem de denetim sağlama bakımından öteki yöntemlere göre daha esnektir,
- ☐ Cevaplayıcı ile işbirliği olanağı vardır,
- ☐ Gözlem yapmaya olanak tanır,
- ☐ Cevaplayıcıdan derinlemesine bilgi edinmek için, ek soru sorulmasına imkan tanır,
- ☐ En üst düzeyde cevaplama oranı sağlar.

Yöntemin dezavantajları ise

- ☐ Görüşmecinin hatalı davranışı görüşmeyi etkiler,
- ☐ Ücretli anketör çalıştırılıyorsa, anketör anketi, hayali bilgilerle doldurabilir,
- ☐ Maliyeti yüksek olabilir,
- ☐ Cevaplayıcıya ulaşmak çok zaman alabilir,
- ☐ Büyük çaplı araştırmalar için yeterli nitelik ve sayıda anketör bulmak zor olabilir,
- ☐ Görüşmeyi yapanın yanlı tutumu denetlenemez.

Telefonla anket

Anket sorularının telefonla cevaplayıcıya sorulduğu ve alınan cevapların anket formuna işlendiği bir veri toplama yöntemidir. Yöntemin avantajlar

- ☐ Ucuzdur, kısa zamanda bilgi sağlar,
- ☐ Denetimi, yüz yüze görüşmeye göre daha kolaydır,
- ☐ Posta ile ankete göre daha esnektir.

Yöntemin dezavantajları ise

- ☐ Herkesin telefonu olmadığı ya da herkesin telefon rehberinde ismi olmadığı için, örnekleme hatalarına düşülebilir,
- ☐ Ayrıntılı bilgi edinme olanağı sınırlıdır,
- ☐ Cevaplayıcı ile kurulacak ilişki yüz yüze görüşmede olduğu kadar sağlıklı olmayabilir,
- ☐ Uzun görüşmelere, şekil ve resim gibi yardımcı araçların kullanımına uygun değildir.

Posta İle Anket

Anketlerin cevaplayıcılara posta yolu ile iletildiği bir veri toplama yöntemidir. Cevaplayıcı anket formunu doldurup araştırmacıya geri gönderir. Yöntemin avantajları

- ☐ Maliyeti, öteki yöntemler göre daha düşüktür,
- ☐ Cevaplayıcı daha bol zamana sahip olduğundan, anketi daha dikkatli cevaplayabilir,
- ☐ Anketör cevaplayıcıyı etkilemez,
- ☐ Sır sayılabilecek bilgiler daha kolay toplanır.

Yöntemin dezavantajları ise

- ☐ Anketin geri dönüş oranı düşüktür,
- ☐ Cevaplayıcıların cevapları denetlenemez,
- ☐ Soruların tam anlaşılıp anlaşılmadıkları denetlenemez,
- ☐ Sorulara kimlerin cevap verdikleri bilinemez.

Internet’te Anket

Anketin bir web sitesinde online doldurulması yada e-posta yoluyla cevaplayıcılara anket formunun gönderilmesi şeklindeki veri toplama yöntemidir. Yöntemin avantajlar:

- ☐ Maliyeti son derece düşüktür,
- ☐ Gerektiğinde resim ve şekil göstermeye elverişlidir,
- ☐ Hızlı bilgi toplar,
- ☐ Anketör kullanılmadığı için anketörün hatalı tutumundan etkilenmez.

Yöntemin dezavantajlar

- ☐ Herkesin internet bağlantısı olmadığından örnekleme hatalarına düşülebilir,
- ☐ Geri dönüş oranı düşüktür,
- ☐ Kimlerin cevap verdiği denetlenemez,
- ☐ Soruların doğru anlaşılıp anlaşılmadığı belli olmaz,
- ☐ Internet kullanımını bilmeyenler cevap veremez,
- ☐ Kimlerin cevapladığı belli olacağından, sır sayılabilecek sorulara cevap verilmek istenmeyebilir.

6. Örnekleme

Bir araştırmada ele alınan problemi çözmek ya da gerekli hipotezleri test etmek için, ihtiyaç duyulan bilgiler ikincil kaynaklardan derlenemiyorsa, bu bilgiler birincil kaynaklardan toplanmak zorundadır. Birincil kaynaklardan bilgi toplamak ise; zaman, maliyet ve diğer nedenlerden ötürü zor bir iş olduğundan, ana kitlenin tümünü gözlem altına alma ya da ana kitleyi tam olarak sayma yerine, ana kitleyi nitelik ve nicelik yönünden temsil eden bir örneğin çekilmesi yoluna gidilir. Böylece örnekten elde edilen bilgilerin, belirli olasılık kademelerinde ana kitle için de geçerli olduğu kabul edilebilir. Bu anlamda örnekleme; ana kitleyi nitelik ve nicelik yönünden temsil edebilecek bir kümenin çekilmesi işlemidir

Örneklemenin amacı tahmin yapmaktır. Tahminde bulunma, örnekleme yoluyla alınacak kısıtlı ve dar bir bilgiye dayanarak daha geniş çaptaki verilerin karakteristikleri, özellikleri hakkında genelleme yapmaktır. Örneğin, bir çuval dolusu pirinç içinden alınacak bir avuç pirincin incelenmesiyle tüm çuval dolusu pirinç hakkında çeşitli özellikleri itibariyle genelleme yapılabilir. Burada, çuvaldaki toplam pirinç miktarı ana kitle (popülasyon)’dır. İncelemek amacıyla aldığımız bir avuç dolusu pirinç ise örnektir

Pozivist yaklaşımın öngördüğü örnekleme, örnekleme kuramına uygun olan örneklemedir. Dolayısıyla, ana kitleyi her bakımdan temsil edecek en doğru örneği elde etmeyi amaçlar. Bu da örnekleme kuramındaki örnekleme süreçlerine tam olarak uymakla sağlanır

Bu anlayışa göre örnekleme süreci şu aşamalardan oluşmaktadır

- ☐ Ana kitlenin tanımlanması,
- ☐ Örnekleme çerçevesinin belirlenmesi,

- ☐ Örnek bireylerinin belirlenmesi,
- ☐ Örnekleme yönteminin tanımlanması,
- ☐ Örnek büyüklüğünün saptanması,
- ☐ Örnekleme planının belirlenmesi,
- ☐ Örneklerin seçimi.

Ana Kitlenin Tanımı

Bir araştırmada, çözüm aranan problemle ya da test edilecek hipotezlerle ilgili sağlıklı ve doğru bilgilerin kimlerden toplanabileceğinin belirlenmesi işlemine ana kitlenin tanımlanması denir

Örnekleme Çerçevesi

Ana kitleden öngörülen miktarda tesadüfi olarak örnek çekilebilmesi için, ana kitleyi kapsayan bir listenin elde bulunması gerekir. Sözelimi, doktorlar ya da avukatlardan bilgi toplanacaksa, tabipler odasından ya da barodan, mesleğe kayıtlı olanların listesi alınabilir. Bir mahalleye ilişkin liste ise, muhtarlıklardan elde edilebilir

Örnek Bireylerinin Belirlenmesi

Örnek bireylerinin belirlenmesinde en kritik nokta, ana kitleyi oluşturan bireylerin örneğe girme ihtimalini eşit hale getirememesi durumunda, sistematik hatanın kaçınılmaz olarak ortaya çıkmasıdır. Örnekleme kuramı, ana kitleyi oluşturan her bireyin örneğe girme şansının eşit olmasını öngörür

Örnekleme Yöntemi

Araştırmanın güvenilirliği bir de, aranan bilginin özelliği gereği örnekleme yönteminin doğru seçilmesine bağlıdır. Bu, aynı zamanda araştırmanın maliyeti ile de ilgilidir. Örnekleme yöntemleri önce tesadüfi ve tesadüfi olmayan olmak üzere iki gruba ayrılır

Örnekleme kuramına göre, örneklerden elde edilen bilgilerin matematik ve istatistik tekniklerle test edilip genelleme yapılabilmesi için örneklemin tesadüfi örneklemeğe uygun olarak yapılması gerekir. Ancak, bu tesadüfi olmayan örneklemin işe yaramaz olduğu anlamına gelmemelidir. Araştırmanın amacına, türüne, maliyetine, kapsamına ve genelleme iddiasına göre tesadüfi olmayan örnekleme yöntemleri de kullanılabilir

Tesadüfi Örnekleme Yöntemleri

Ana kitle hakkında tahminde bulunurken yapılan tahminin geçerli olabilmesi için örneklemin tesadüfi (rastgele) olması gerekir. Tesadüfi örnekleme kavramı, örnekte yer alacak elemanların belirlenmesinde hiçbir dış etkinin rolünün olmamasını ifade eder. Diğer bir ifadeyle, bir ana kitleden herhangi bir örnek alan bir kişi, örnekte bulunsun veya bulunmasın hiçbir eleman veya birim üzerinde etkiye sahip olmayacaktır

Basit Tesadüfi Örnekleme

Bu yöntemin kullanılabilmesi için, ele alınan sorun ya da hipotezlerle ilgili bilgilerin ana kitleye göre homojen (türdeş) olması gerekir

Basit tesadüfi örneklemede, ana kitledeki her birimin örneğe girme olasılığını eşit hale getirmek esastır. Bunu sağlamak için değişik yollara başvurulabilir. Ana kitle listesi bilgisayara yüklenerek buradan rastgele çekiliş yapılabilir. Bireylere birer numara verilerek bu numaralar bir torbaya konarak rastgele bu numaralar çekilebilir Rastgele sayılar tablosu kullanılarak da örneğe alınacak birimler seçilebilir.

Yöntemin Yararlı Yönleri

- ☐ Evrendeki her elemanın eşit seçilme şansı vardır,
- ☐ Evren çok büyük ve karmaşık değilse seçme işlemi kolaydır,
- ☐ Bu yöntemle yapılan örneklemede istatistiksel işlemler ağırlıksız olarak yapıldığı için değerlendirme işleminde kolay olur.

Yöntemin Sakıncalı Yönleri

- ☐ Evren çok büyükse evreni listelemek ve seçmek güçtür,
- ☐ İncelenen özellik evrendeki elemanların bazı özelliklerine göre değişiklik gösterebilir,
- ☐ Örnekleme seçilecek bireyler çok geniş bir bölgede dağınık bir şekilde yerleşmiş olabilirler,
- ☐ Evrende azınlıkta olan grupların yeterince temsil edilmemesi söz konusu olabilir.

Zümrelere Göre Örnekleme (Tabakalı Örnekleme)

Elde edilecek bilgi, ana kitle itibari ile türdeş değilse yani, bilgi ana kitleyi oluşturan değişik özellikteki gruplara göre farklılık gösteriyorsa, zümrelere göre örnekleme yöntemi kullanılır. Ana kitle zümrelere (tabakalara) ayrılırken dikkat edilmesi gereken husus, elde edilen bilgi ya da ele alınan sorun hangi kriterlere göre farklılık gösteriyorsa, zümrelerin de bu kriterlere göre oluşturulmasıdır. Sözelimi, ele alınan sorun ya da sorunu çözecek bilgiler sosyo-ekonomik ya da sosyo-kültürel faktörlere göre farklı ise, zümreler de bu faktörlere göre oluşturulmalıdır Zümreler belirlendikten sonra,

her zümreden belirli sayıda birim rastgele örnekleme ile çekilmelidir.

Her tabakaya eşit sayıda birey düşmesi olanaksız olacağından, her tabakadan kaç bireyin örnekleme alınacağı sorunu çıkar. Bu durumda iki yol izlenebilir. Birincisinde, tabakalardaki birey sayısı göz önüne alınmadan her tabakadan eşit sayıda birey örnekleme alınır. Buna orantısız seçim denir. Orantısız seçimde istatistiksel değerlendirmenin kesinlikle ağırlıklı olarak yapılması gerekir. İkincisinde ise, örnekleme alınacak bireyleri tabakalardaki birey sayısına orantılı olarak seçmektir. Başka bir deyişle, çok kişi içeren tabakadan çok, az kişi içeren tabakadan az kişiyi örnekleme almaktır.

Örnekleme seçimi orantılı yapıldığında aritmetik ortalama ağırlıksız, standart sapma ise ağırlıklı olarak hesaplanır. Orantılı seçim, işlemleri kolaylaştırdığı için tercih edilen bir yoldur.

Tabakalı rastgele örnekleme yöntemine tabakalar arasında gerçek bir farklılık olduğunda başvurulmalıdır. Bu yöntemin sakıncalı yanları çok azdır. Bunlar;

- ☐ Tabakalardaki birey sayısının bilinmediği durumlarda seçim işlemlerinin güçleşmesi,
- ☐ Örnekleme seçilecek birimlerin çok büyük bir bölgede dağınık olarak oturması durumunda araştırmanın uygulama aşamasının güçleşmesidir.

Kümelere Göre Örnekleme

Bu örnekleme yönteminde, en güvenilir örneği elde etme yerine, en düşük maliyetle en doğru örneği elde etmek amaçlanır. Ana kitle önce kümelere ayrılır, sonra kümelerden bireylere geçilir

Bu tür örnekleme, sağlıklı bir ana kitle çerçevesinin elde bulunmaması ya da çok büyük ana kitleden çekim yapmanın çok zor ve maliyetinin yüksek olması durumunda uygulanır. Bu tür örnekleme tesadüfi örnekleme olmasına karşın, ana kitleyi ne ölçüde temsil ettiği tartışılabilir. Bu nedenle de çok dikkatli yapılması gerekir

Küme örnekleme yönteminde genel kural kümedeki birim sayısının az olması yani kümelerin küçük olmasıdır. Kümelerin küçük olması küme sayısını artıracak, bu da değişik özellikteki kümelerin örnekleme girme şansını arttıracaktır. Örneğin 5 000 aile içeren bir bölgeyi 1000'er ailelik 5 kümeye ayırıp buradan 1 kümeyi örnekleme alma yerine, 250'şer ailelik 20 kümeye ayırıp 4 küme seçmek daha uygundur.

SistematiK Örnekleme

Örnekleme seçim işlemlerinin kolay olması nedeniyle özellikle evren büyük olduğunda kullanılan bir örnekleme yöntemidir. Bu yöntemin en çok kullanıldığı durumlar:

- ☐ Çok sayıda birim içeren kayıt sistemlerinin incelenmesinde. Örneğin, hasta dosyaları, hasta ya da işçi kayıtları, kayıt defterleri, fişler , listeler gibi.
- ☐ Birim sayısı çok fazla olduğu için listelenmesi güç ya da olanaksız olan durumlarda. Örneğin, büyük bir kentte ev seçimi, sokak seçimi, işyeri seçimi otomobil seçimi gibi.

Seçim işlemlerinde evren büyüklüğü (N) örneklem büyüklüğüne (n) bölünerek kaç birimde bir birimin örnekleme alınacağı saptanır. Örneğin, 15 000 hasta dosyası bulunan bir arşivden 500 dosya örnekleme seçilecekse ($15\ 000 / 500 = 30$) her 30 dosyada bir dosya örnekleme alınacaktır. Başlangıç sayısı rastgele sayılar tablosundan 1 – 30 arasında bir sayı seçilerek bulunur. Seçilen sayı 8 ise önce 8'inci dosya örnekleme alınır, sonra her 30 dosya 1 dosya örnekleme alınır. Böylece örnekleme çıkan dosya numaraları 8, 38, 68, 98,14 978 olacaktır.

Bu yöntemi kullanacak araştırmacılar şu noktaları göz önünde bulundurmalıdırlar

- ☐ Başlangıç sayısı dağılımı büyük oranda etkiler.Örneğin dosyalar küçük yaştan büyük yaşa doğru sıralanmışsa ve araştırmacı yaş ortalamasını öğrenmek istiyorsa 3.33.63.93.... sırasında elde edilecek ortalama ile 28,58,88,118.... Sırasından elde edilecek ortalama farklı sonuçlar vermektedir.

- Birden çok kurumda dosyalar incelenecekse ve her kurumda diyelim 30 dosya varsa ve her kurum dosyaları küçük yaştan büyük yaşa doğru sıralanmış ise başlangıç sayısı dağılımı yine etkiler.
- Birden çok kurumda dosyalar incelenecekse bir kurum dosyaları büyükten küçüğe bir kurum dosyaları küçükten büyüğe doğru sıralanmışsa bir diğeri de sırasız olarak dizilmişse araştırmacı bunların sırasını belirli bir düzene soktuktan sonra seçim işlemine geçmelidir.

Tesadüfi Olmayan Örneklem Yöntemleri

Tesadüfi örneklemenin uygulanamadığı durumlarda uygulanan ölçekleme yöntemlerine verilen addır. Nicel araştırmalarda kullanılmaları, örneklem kuramının genelleyici özelliğine uymadığından, kullanılmaları eleştirilir. Buna karşılık nitel araştırmacıların ve ön araştırma yapanların sıkça bu yöntemleri kullandıkları görülmektedir.

Kolay Yoldan Örneklem (Kolayda Örneklem)

Bu tür örneklem en düşük maliyetli ve uygulanması en kolay örneklemedir. Süpermarket çıkışındaki tüketicilere, sokakta dolaşan insanlara uygulanan anketler buna örnek gösterilebilir.

Yargısal Örneklem

Ana kitleyi temsil edeceği varsayımı ile yapılan örneklemedir. Yeterince deneyimlilerin ve uzman kişilerin görüşleriyle hangi örneklerin ana kitleyi temsil edebileceği belirlenir.

Kota Örneklemesi

Kota örneklemesinin uygulanabilmesi için, toplanacak bilginin hangi seçim kriterine göre değerlendirileceğine karar vermek gerekir. Örneğin, toplanacak bilgiler yaş, cinsiyet ya da öğrenim düzeylerine göre farklılık gösteriyorsa, bunların ana kitle dağılımına uygun örnek dağılımı elde edilmelidir. Sözelimi, yüksek öğretimli olanların ana kitle içindeki oranı % 10 ise, örnek içindeki dağılımı da %10 olmalıdır.

Kartopu Örneklemesi

Bu örneklemede, araştırmacı ilk adımda tamamen rastlantısal olarak seçtiği çekirdek bir örnekle yola çıkar. İkinci adımda, çekirdekte yer almış örneklerin önerileri ile yeni örneklere ulaşır ve işleme böylece devam edilir. Sözelimi, matematik öğretimi ile ilgili bir araştırma yapacaksa, ilk adımda bulabildiği üç beş matematik öğretmeni ile görüşür ve daha sonra onlardan başkaca önerdikleri matematik öğretmenlerinin isim ve adreslerini alarak onlarla görüşür.

Örnek Büyüklüğünün Belirlenmesi

Bazı durumlarda, yapılan tahminlerde belli bir doğruluk veya isabet derecesine ulaşmak hedef olarak alınabilir. Bu gibi durumlarda istenen hedefe ulaşabilmek için ne büyüklükte örnek alınması gerektiği araştırma konusudur.

Örnek büyüklüğünün hesaplanmasında aşağıdaki formüller kullanılabilir.

$$n = \frac{Z^2 \sigma^2}{(\bar{x} - \mu)^2} \left\} \begin{array}{l} \text{Ana kitle varyansı biliniyorsa} \\ \text{Ana kitle varyansı bilinmiyorsa} \end{array} \right.$$

Z : belirlenen güven düzeyi için Z tablosundan bakılan değer,

σ^2 : ana kitle varyansı,

$\bar{x}$: örnek ortalaması,

μ : ana kitle ortalaması,

$(\bar{x} - \mu)$: göze alınan örnekleme hatası.

Örnek : Bir iş kolunda çalışan işçilerin ortalama saat ücretleri gerçek ana kitle ortalama saat ücretinden en fazla 100 TL sapma gösterecek şekilde ve %90 güven aralığında tahmin edilmek isteniyor. Geçmiş kayıtlara dayanarak bu iş kolu için hesaplanan standart sapma olduğuna göre örnek büyüklüğü ne olmalıdır.

Bu sonuca göre, gerçek ortalamanın 100 TL sınırları içinde ve % 90 güven aralığında istenen amacı gerçekleştirmek için örnek büyüklüğü 68 olmalıdır.

Örnekleme Planının Belirlenmesi

Örnek büyüklüğü de belirlendikten sonra örnekleme planı oluşturulur. Bu planda hangi bilgilerin kimler tarafından, ne zaman kimlerden toplanacağı ve örnekleme için ayrılan bütçe belirlenir. Araştırmada görev yapan herkes toplayacağı bilgileri, bilgi toplama yerlerini, hangi bilgileri toplayacaklarını, bilgi toplayacağı kişileri ve kullanabilecekleri kaynakları bu planda görebilmelidir. Araştırma bu plan çerçevesinde gerçekleştirilir.

Örnekleme Hataları

Örnekleme yöntemlerine dayalı olarak yapılan tahminlerde iki tip hata ile karşılaşılır. Birinci tip hata, tesadüfi hata olarak adlandırılır. Bu tür hatalar örnek sayısı arttırılarak giderilebilir. Sistematiik hata olarak adlandırılan ikinci tip hatalar ise; örnekleme sürecindeki hatalardan kaynaklanır ve sonradan giderilmeleri mümkün değildir. Bu hatalar:

- ☐ Örnekleme yönteminin doğru seçilmeyişinden,
- ☐ Ana kitlenin yanlış tanımlanmasından,
- ☐ Örnekleme çerçevesinin yanlış belirlenmesinden,
- ☐ Örneklerin doğru çekilmeyişinden,
- ☐ Örnek büyüklüğünün doğru hesaplanmayışından kaynaklanırlar.