

ÖRNEKLEMESİNE DAYALI ARAŞTIRMALARIN TARİHÇESİ

İstem Köymen KESER*

ÖZET

Örnelemeye dayalı araştırmaların geçmişi tarım, kamu yönetimi ve sosyal araştırmalar gibi birçok alanda çeşitli faaliyetlere dayanmaktadır. Kamu yönetimi ve sosyal araştırmalarda hükümetler ve bireysel araştırmacılar tarafından büyük veri setlerinin toplanmasının rastsal örnekleme yöntemlerinin gelişimine büyük bir katkısı olmuştur. Bu çalışmada da örnelemeye dayalı araştırmaların bizi rastsal örnekleme yöntemlerine götüren tarihsel gelişim süreci özetlenmektedir. Örnekleme teorisi ve uygulamaları bugün modern istatistik biliminde ayrı ve büyük bir dal haline gelmesine rağmen başlarda durum kesinlikle bu değildir. Örnelemeye dayalı araştırmaların kökleri olasılık teorisi ve deney tasarımıydansa daha çok resmi istatistik ve sosyal istatistiklere dayanmaktadır. Özellikle politik aritmetik örnekleme modern örnekleme teorisine götüren önemli öncül faaliyetlerdir. Bununla birlikte olasılık teorisi örnekleme teorisinin önemli bir bileşeni haline geldikten sonra örnelemeye istatistik biliminin önemli bir dalı olarak bakılmaya başlanmıştır. Bu çalışmada sosyal istatistiklerin elde edilmesinden günümüz modern örnekleme teorisine kadar olan süreç, özellikle örnekleme araştırmalarının gelişimi için büyük önem taşıyan 1895-1934 dönemi ve Kiaer, Bowley ve Neyman dönüm noktaları detaylı olarak incelenmiştir.

Anahtar Kelimeler: Amaca yönelik örnekleme, Örnelemeye dayalı araştırmalar, Rastsal örnekleme, Tarihsel gelişim, Temsil edici metod.

1. GİRİŞ

Bilimde örnekleme prensiplerinin sezgisel uygulamaları uzun zamandır yer almaktadır, ancak başlarda örnekleme olarak değil de tümevarım olarak adlandırılmaktadır. Birçok bilimsel sonuç sadece birkaç deneydeki gözlemlere dayanır. Tümevarım hem günlük yaşamda ve hem bilimde uzun zamandır uygulanmasına rağmen, iyi tanımlanmış bir istatistiksel metod olarak örnekleme oldukça gençtir. Çoğu bilim adamına göre resmi olarak başlangıç tarihi 19. yüzyılın sonlarında 1895 yılı olarak kabul edilmektedir (Bethlehem, 2009).

Örnelemeye dayalı araştırmalar yüzyıllardır sezgisel olarak başlamasına rağmen yaklaşık 1900'lere kadar örnekleme ile ilgili bir bilimsel teori gelişmemiştir. Ancak 1900'lerden önceki dönemlerde de tam bir nüfus sayımının yapılamadığı durumlarda kitle hacminin tahmin edilmesi ile ilgili konular birçok bilim adamının aklını meşgul etmiştir. Bir ülkenin nüfusunun yüksek olması o dönemlerde özellikle diğer ülkelere göre güç bakımından önemli bir göstergedir. Ancak her zaman nüfusun tam bir sayımının yapılması mümkün değildir. Önceleri belki de öncelikli olarak güç simgesi olarak görülen nüfusun tahmin edilmesi amacıyla başlayan o zamanlar kısmi araştırma olarak adlandırılan örnelemeye dayalı araştırmalar daha sonra sosyal, ekonomik ve tarımsal v.b. alanlarda da tam sayıma göre daha detaylı bilgi toplayabilmek amacıyla kullanılmaya başlanmıştır.

*Yrd. Doç. Dr., Dokuz Eylül Üniversitesi İ.İ.B.F., Ekonometri Bölümü, İstatistik Anabilim Dalı, e-posta: istem.koymen@deu.edu.tr

Özetle, modern örnekleme teorisine giden süreç birkaç temel dönemden geçmiştir ve bu dönemler bulgular başlığı altında ayrıntılı olarak incelenmiştir. Bu temel dönemler ayrıca incelenen örnekleme yöntemlerinin özellikleri nedeniyle temsil edici örnekleme, amaca yönelik örnekleme ve son olarak da bugünkü olasılıksal örnekleme yöntemlerinin temelini oluşturan rastsal örnekleme olarak da sınıflandırılabilir.

Bu çalışmanın amacı modern örnekleme teorisine temel olan özellikle 1895’de başlayıp 1934 rastsal örnekleme kadar olan süreci etkileriyle birlikte dönemsel bazda ele almaktır. Burada örnekleme yöntemlerinin daha çok tarihsel gelişimleri, o dönemde yaşanan tartışmalar, yapılan katkılar, gelişimler, örnekleme yöntemlerinin nasıl uygulandığı ve rastsal örnekleme yöntemlerine giden süreç ile ilgili detaylar üzerinde durulmuştur. İlk defa kısmi bilgiden yararlanılmaya başlanılan 1662’lerde demografinin kurucusu olarak adlandırılan John Graunt ile başlayan öncül gelişimlerin de örneklemeyle dayalı araştırmaların başlamasına itici güç olduğu yapılan incelemelerden anlaşılmaktadır. Teorik incelemeler çalışmanın kapsamı dışındadır.

2. YÖNTEM

Bu çalışmada bugünkü modern örnekleme teorisine ve uygulama yöntemlerine temel oluşturan örneklemeyle dayalı araştırmaların tarihçesi dönemsel bazda incelenmiştir. Öncelikle 1895 dönemi öncesi özellikle devletin ihtiyaçları doğrultusunda bilgi edinilmek amaçlı kullanılan kısmi araştırmalar incelenmiştir. Bu araştırmalar tam bir sayımın yapılmasının mümkün olmadığı durumlarda ortaya çıkan kısmi araştırmalara duyulan ihtiyacı göstermektedir. Daha sonra ilk defa “temsil edici örnek” terimi kullanılarak yeni bir dönem başlamasına neden olan Kiaer ve damgasını vurduğu 1895–1903 periyodu 1895–1924 dönemi altında incelenmiştir. Daha sonra 1924–1934 dönemi altında, Uluslararası İstatistik Enstitüsünün (International Statistical Institute (ISI)) özellikle örneklemeyle dayalı araştırmaları çalışması için oluşturduğu komisyonun çalışmalarına ve bu komisyondaki temel güçlerden biri olan Bowley’in örnekleme konusundaki 1895–1924 döneminde başlayan katkılarının devamına değinilmiştir. Son olarak da 1934’de Neyman ile örnekleme de olasılıksal yöntemlerin de açık bir biçimde devreye girmesiyle yeni bir dönem başlamıştır. Bu dönem ve sonrası da ayrı bir başlık altında ele alınmıştır.

Bu çalışmada çok sayıda makale incelenmiş ve konu ile ilgili güvenilir en eski kaynaklara da ulaşılmaya çalışılmıştır.

3. BULGULAR

3.1 Örneklemeyle Dayalı Araştırmalarda Bazı Öncül Gelişimler: 1895 Dönemi Öncesi

Kitleden bir kesim olarak kitle hakkında yorumlamalarda bulunmak için bilinen ilk girişimlerden biri John Graunt (1620–1674)’ındır. Bethlehem (2009)’da belirtildiği gibi Graunt (1662) kısmi bilgiye dayanarak Londra’nın nüfusunu tahmin etmek için bir yöntem önermiştir. Graunt kiliselerin tuttuğu ölüm kayıtlarından yararlanmıştır. Bu tarihte kiliselerin yaklaşık 100 yıllık kayıtları bulunmaktadır. Graunt kayıtların iyi saklandığı papazların yönetimindeki bölgelerden seçilen bir örnekteki aileleri incelemiştir. 11 ailede yıl başına ortalama olarak 3 ölümün olduğunu bulmuştur. Bu oranın tüm papaz yönetimindeki bölgelerde yaklaşık olarak eşit olduğunu varsaymıştır.

Londra'daki toplam ölüm sayılarının yıl başına yaklaşık 13,000 olduğunu bilmektedir. Buradan ailelerin toplam sayısının 48,000 olduğu sonucunu çıkarmıştır. Ortalama aile büyüklüğü 8 olarak alındığında Londra'nın nüfusu 384,000 olarak tahmin edilmiştir.

Graunt aile başına ölüm sayısı gibi ortalamaların bir aralıkta ve zamanda değiştiğinin farkında olmasına rağmen bununla ilgili herhangi bir koşul koymamıştır. Bu yöntemi uygun bir bilimsel temele oturtamamıştır.

Graunt Londra'nın sadece kitle büyüklüğünü tahmin etmekle kalmamış aynı zamanda ölüm kayıtlarında kadın-erkek ve kentsel-kırsal ölüm oranlarını da kıyaslamak gibi farklı demografik çalışmalarda uygulamıştır. Bu nedenle John Graunt'dan sıklıkla demografinin (nüfus istatistikleri bilimi) kurucusu olarak da bahsedilir.

Ayrıca 17. yüzyılda öncelikle İngiltere'de ve bazı diğer Avrupa ülkelerinde özellikle Graunt ve Eden tarafından 1650 ve 1800 yılları arasında uygulanan bu hesaplamalar politik aritmetik olarak adlandırılır ve politik aritmetik tartışmaların önemli bir kısmı haline gelmiştir. Politik aritmetik oransal tahminlere (ratio estimate) benzer fikirlere dayanır. Seçilen bölgelerdeki doğum oranları, aile hacimleri, evlerdeki ortalama kişi sayısı ve bilim adamlarının kişisel gözlemlerine bakılarak kitle hacmini tahmin etmek mümkündür. Bu tahminlerin bazıları daha sonra bazı nüfus sayımlarınca doğrulanmıştır. Benzer gelişimler Fransa ve Belçika'da da uygulanmıştır (Biemer ve Lyberg, 2003). Özellikle William Petty (1623 – 1687) de politik aritmetik konusunda birçok çalışma yapmıştır. Bu çalışmaların üzerine hükümet Petty'yi İrlanda'nın derinlemesine bir araştırmasını yapması için atamış ve Petty'de İrlanda'nın etkileyici bir analizini yapmıştır (Brunt, 2001).

1600'lerde yine örnekleme dayanılarak yapılan bir çalışma Halley kuyruklu yıldızının yörüngesini ilk kez hesaplamış olan İngiliz matematikçi ve astronomi bilgini Edmond Halley (1656–1742) tarafından 1693 yılında oluşturulmuştur. 1693'te nüfus kayıtlarının düzenliliğiyle tanınan Alman-Polonya kenti Breslau'nun ölüm yaşı kayıtlarını analiz ettiği bir makale yayınlamıştır. Bir şehirdeki bu örneğe dayanarak insanoglunun ölüm oranı ile ilgili genellemeler yapmıştır (Stephan, 1948). Bu çalışmada Breslau'nun 1687–1691 yılları arasında 5 yıllık dönem boyunca demografik verileri incelenmiştir. Bu makale, İngiliz hükümetine hayat sigortası fiyatlarını alıcının yaşına bağlı olarak ayarlama imkânı vermiştir. Halley'in çalışmaları aktüeryanın gelişimine büyük katkıda bulunmuştur (Bacaer, 2011).

1754'lerin başında İngiltere'nin nüfus tahminleri vergi listesindeki evlerin sayıları ve vergilendirilmemiş küçük evlerin kaba tahminlerine göre yapılmaktadır. Konutların toplamı biraz keyfi bir sabit olan 6 kişi ile çarpılmakta ve diğer tahminler kayıtlı vaftiz, evlilik ve cenaze töreni sayılarına göre yapılmaktadır (Stephan, 1948).

Örnekleme genelde resmi sayımların oluşturulmasından önce gelmektedir. 1801'deki İngiltere'deki sayımlardan önce İngiltere'nin nüfusu tam olarak doğru bir biçimde bilinmemektedir ve bu konuda çeşitli spekülasyonlar yapılmış, tahminler hazırlanmış ve tartışılmıştır. Periyodik ulusal sayımların oluşturulmasından önce karamsar kuramlarıyla ünlenen ünlü nüfus bilimci ve ekonomi politik teorisyeni Thomas Robert Malthus (1766 - 1834) 1798 yılında, nüfusbilimi için çok önemli kurallara imza atan ve tanınmasına neden olan "Nüfus Hakkında Bir Deneme" (An Essay On Population) adlı çalışmasını yayınlamıştır. Çalışması büyük yankılar uyandırmış ve birçok yeni

tartışmaya neden olmuştur. Çalışmasına göre uygun şartlarda herhangi bir nüfus, besin maddelerinin artışından daha hızlı bir oranda artar ve böylece zamanla kişi başına düşen besin miktarının azalacağını ve kitlesel bir açlığa yol açacağını öngörmüştür. Bu fikrinin temelinde ise uygun şartlarda herhangi bir kısıtlayıcı faktör (salgın hastalıklar, ekonomik kriz, savaş v.b.) yoksa nüfusun geometrik dizi biçiminde oysa besin maddelerinin aritmetik dizi biçiminde artacağı yatmaktadır (Malthus, 1798; Stephan, 1948).

Yine bu dönemlerde Arthur Young (1741–1820) 1765 ve 1815 yılları arasında yarım yüzyıllık bir dönemde tarımsal ve ekonomik eleştirmenlerin liderlerinden biridir. Young o zamanlar İngiliz tarımının tam bir sayımını yapmak için yeterli kişisel kaynaklara ve otoriteye sahip değildir ve bu sebeple Young gereksinim duyulan verileri sağlamak için yeni bir fikir önermiştir. Bu fikir çiftliklerden oluşan bir örneğin incelenmesidir ve bu amaçla 1768’de İngiltere’de örneklemeye dayalı kırsal ekonomi ile ilgili geniş çaplı bir araştırma yapmıştır. Bu araştırmanın vurgulanan üç önemli noktası vardır. Birincisi araştırmacı bizzat verileri kendi topladığından daha güvenilir olacaktır. Ayrıca 19 yy.da tarımsal sayımları yapan istatistikçiler çiftçilerin yeterli zamanı olmadığına ve arşivciliklerinin zayıf olduğuna inanmaktadırlar. İkincisi araştırmacı özellikle ücra bölgelerin içerisinden de geçerek temsil ediciliği arttıracaktır. Üçüncü olarak da daha çok değişkenle ilgili veri toplanabilecektir. Özetle tamsayım olmadan konuyu aydınlatmak ve kullanışlı veri toplamak mümkün hale gelebilecektir. Bu çalışma 413 çiftliği ve tüm çiftlik türlerini kapsamaktadır, ayrıca 400 değişkene ait veri toplanmıştır (Brunt, 2001).

1800’de Sir Frederick Morton Eden (1766–1809) Büyük Britanya’nın nüfusunu 9,000,000 olarak tahmin etmiştir. Burada örneğini ev başına ortalama hane halkı ve doğum sayısına dayandırmıştır. Daha sonra Büyük Britanya’da 1801’de yapılan ilk sayım bu tahmini doğrulamıştır (Stephan, 1948).

1864’de büyük reformlar doğrultusunda Rusya’da Zemstvo’lar (yerel meclisler) kurulmuştur. Bunlar yerel yönetimleri düzenleyen ve üyeleri seçimle belirlenen meclislerdir. Zemstvo’lar kendi kendilerine istatistik bölümleri aracılığıyla Rusya’nın kırsal kesimiyle ilgili araştırmaları başlatmışlardır. Rusya’da da diğer Avrupa ülkelerinde olduğu gibi örneklemeye dayalı ilk araştırmalar belirli bir bölgenin her tarafına ulaşılmasının mümkün olmadığı ve tam sayım yapılamadığı durumlarda ekonomik ve sosyal ölçümleri toplamak için sayısal bilgi toplama ihtiyacından doğmuştur. Bu ihtiyaca yönelik ilk girişimler 1870’lerde başlamıştır, yönetsel bölgenin sadece bir kısmına bir kısmi araştırma uygulamışlardır. Rus İstatistikçiler örneğe “tipik” olarak adlandırılan köyleri seçmişlerdir. Bu köyler bütün köylerle karşılaştırıldığında karakteristikleri ortalama olarak ele alınabildiğinden dolayı tipik olarak adlandırılır. Tipik olarak ele alınan bir köyün çevre bölgede bulunan köylerle karakteristiklerinin çoğu ortaktır. Ancak o zamanlar tipik bölge seçmede kullanılan yöntem daha çok materyal kısıtlarının bir sonucudur. Daha sonra 1876 ve 1895 arasında da yönetim ihtiyaçlarından dolayı bazı örneklemeye dayalı araştırmalar uygulanmıştır. Ancak bunlar tamsayım karşı bir meydan okuma değildir. Ancak yine de örneklemeye dayalı araştırmaların öneminin farkına varılmaya başlanmıştır. Avrupa’da olduğu gibi Rusya’da da temsil edicilik (representativeness) kavramı 1895–1924 periyodu sırasında meydana gelmiştir (Seneta, 1985; Mespoulet, 2002).

3.2 1895–1924 Dönemi: Kiaer ve Bowley Dönüm Noktaları

Yukarıda bahsedilen örnekleme dayalı öncül araştırmalara rağmen 19. yüzyıl boyunca hükümetlerin istatistik kurumlarında ve sosyal reformcular arasında genel kabul gören yöntem tamsayıdır. Örnekleme dayalı yapılan araştırmalar bir bilimsel temele oturtulamamıştır ve bunlar kuraldansa istisna olarak kabul görmüştür. Örneklemin bilimsel olarak başlangıç tarihi çoğu bilim adamı tarafından 1895 yılı olarak kabul görmektedir. Tamsayımın pozisyonu Norveçli istatistikçi Anders Nicolai Kiaer'in (1838–1919) girişimleri ile değişmiştir (Bellhouse, 1988). Kiaer'den bugün resmi istatistiklerde ve sosyal araştırmalarda geniş ölçüde örnekleme dayalı araştırmaların kurucusu ve savunucusu olarak bahsedilmektedir. Kiaer Norveç Merkez İstatistik Bürosunun (Norwegian Central Bureau of Statistics) başkanıdır ve 1876'dan 1913'e kadar bu görevi yürütmüştür. İstatistik Bürosunun başı olarak kitlenin on yılda bir olan sayımlarından, yüzyılın son çeyreğindeki tarım endüstrisinden, nüfus hareketinin ölçümünden, birçok büyük ölçekli istatistiksel araştırmadan o sorumlu olmuştur. Kiaer'in tamsayımdan bağımsız olarak veri toplamada örnekleme metodunu kullanan ilk kişi olduğu söylenebilir. Sistematik olarak bu metodun kullanımını için durumlar geliştirmiş ve başında olduğu büro aracılığıyla çeşitli örnekleme dayalı araştırmalar uygulamıştır (Seng, 1951).

Kiaer 1895'de Uluslararası İstatistik Enstitüsünün (International Statistical Institute (ISI)) Berne'deki toplantısında, "Temsil Edici Metot" (Representative Method) olarak adlandırılan bir metoda dayanan bir kısmi araştırma ile kullanışlı ve tamsayıma göre daha ayrıntılı bilgi sağlanabileceğini öne sürmüştür. Kiaer'in yaklaşımına göre bir kısmi araştırma olan örneğe dayanan araştırmalarda örnek kitlenin yaklaşık bir minyatürü olmalıdır (Bellhouse, 1988; Desrosieres, 2002). Kiaer bu toplantıda 1891 sayımına dayanan bir araştırmadan elde ettiği sonuçları sunmuştur. Bu araştırmanın amacı Norveç'teki yetişkin erkeklerin gelirlerinin yaşa, sivil statüye ve mesleğe göre dağılımını çıkarmaktır. Kullanılan tasarım bugün üç aşamalı tasarım olarak adlandırılmaktadır. İlk aşamada ülke çapında 127 temsil edici bölgenin ve 23 temsil edici kasabanın ve ikinci aşamada da tüm geri dönüşlerden yaşları 17, 22, 27, 32...97'ye kadar olan erkeklerden oluşan bir örnek seçilmiştir. Araştırmada böylece her bir temsil edici şehir veya kırsal bölgedeki erkeklerin sadece beşte biri seçilmiştir. Son olarak üçüncü aşamada isimleri A, B, C, L, M, N harfleriyle başlayan erkekler seçilmiştir. Çerçeve olarak da 1891 sayımı kullanılmıştır. Örneğin temsil ediciliğini değerlendirmek üzere Kiaer araştırmadan elde edilen birkaç marjinal tabloyu 1891 sayım tablolarının sonuçlarıyla karşılaştırmıştır. İncelendiğinde tamsayım ve örnekleme dayalı araştırma sonuçlarına ilişkin dağılımlar olumlu temsiller sağlar ancak bazı noktalarda farklılıklar bulunmaktadır. (Seng, 1951; Beaud ve Prevost, 1998; Thomsen ve Zhang, 2012).

Kiaer'in bilimsel anlayışı yalnızca tamamen rastsal (random) seçimin bir örnek seçmek için tek ve en iyi metot olmadığını farkındaydı. Uluslararası İstatistik Enstitüsünün Berne'deki toplantısında sunduğu raporunda da ülkenin koşullarını çalışmak için ülke boyunca küçük coğrafi birimlerden veya yerleşimlerden yeterli sayıda örnek alınması gerektiğini belirtmiştir. Eğer birimlerin seçimi rasyonel bir biçimde yapılırsa birimler ülkeyi temsil edebilirler. Dolayısıyla başarılı bir örnekleme için iki önemli koşuldan bahsedilebilir: uygun temsil (proper representation) ve birimlerin rasyonel seçimi (Seng, 1951).

Uygun temsil için kullanılan metot bir çeşit tabakalamadır. Tabakalar coğrafi, sosyal ve ekonomik olabilir. Kiaer'in yukarıda bahsedilen çalışmasında da coğrafi tabakalama yapılmıştır. Coğrafi tabakalama, evren kasabalara ve kırsal bölgelere ayrılarak başarılmıştır. Daha sonra kasabalarda sosyal koşullarla ilgili tabakalama için caddeler zengin, orta ve fakir kesimleri gösterecek biçimde kategorilere ayrılmıştır (Seng, 1951). Kırsal bölgelerde de temel uğraşlara (occupations-meslek) göre bir tabakalama yapılmıştır. Ayrıca görüşmeyi yapacak olan kişilere görüşme için sadece sosyal olarak ortalama olan evleri değil farklı sosyal ve ekonomik koşulları temsil eden evleri de seçmeleri istenmiştir. Tabii ki insan seçiminden kaynaklanan sapma kaçınılmazdır fakat en azından tabakalama metodu açıktır. Ayrıca Kiaer her bir seçilen tabakada önceki sayımın kitle detaylarına dayanarak orantısal seçimi (proportional selection) önermiştir (Seng, 1951; Statprob, 2011). Tabakalamanın eksiksiz teorisi Neyman'ın zamanına kadar geliştirilmemiştir, ancak yine de Kiaer çoklu faktör tabakalamaya gereksinim duyan bir örnekleme tasarımını geliştirmek için bir kullanışlı modelin taslağını oluşturmuştur (Seng, 1951).

Kiaer'in temsil edici örnekle ne kastettiğini Kruskal ve Mosteller (1980) şu şekilde özetlemişlerdir: "Kiaer ilk olarak sosyal ve ekonomik araştırmaları ele almıştır. Burada evler, aileler, bireyler gibi birimlerin seçiminde ilk olarak bölgeler, kasabalar, şehirlerin bazı kısımları, caddelerin seçiminde v.b. sistematik bir yol izlenecektir. İkinci olarak böyle bir seçim sürecinin tüm seviyelerinde tatmin edici örnek hacimlerinde ısrar etmiştir. Üçüncü olarak örneği çeşitli yollarla özellikle coğrafi olarak yayma ihtiyacını vurgulamıştır".

Sıklıkla çeşitli modlarda kapsam fikrine geri dönmüştür. Örneğin şöyle demiştir "Sadece sosyal açıdan bakıldığında ortalama olan aileleri değil, toplumda farklı ekonomik ve sosyal koşulları olan aileleri de ziyaret etmeye dikkat etmeliyiz" (Kruskal ve Mosteller, 1980).

Kiaer kendi temsil edici örneğini kitlenin yaklaşık bir minyatürü olarak düşünmüştür. Bu fikri zaman zaman tekrarlamış ve örnek sonuçlarının her nerede mümkünse, örneğin yaş, medeni durum, iş v.b. açısından, bilinen tam sayım sonuçları ile karşılaştırılmasının önemini vurgulamıştır. Tamsayıma nazaran örnek almanın avantajı çok daha fazla değişken ile ilgili bilgi sağlayabilir. Bu bir tamsayım için oldukça maliyetli olabilir (Kruskal ve Mosteller, 1980).

Kiaer araştırmancının bilimsel değerinin gözlem sayısından daha çok temsil edici karakterlere bağlı olduğunu vurgulamıştır. Ancak yine de örnek seçiminin tüm seviyelerinde tatmin edici örnek hacminde ısrar etmiştir. Kiaer temsil ediciliği elde etmenin çeşitli yolları olduğunu iddia etmiştir. Eğer karşılaştırmalar kısmi araştırmancının sonuçları ile tamsayımın sonuçları arasında benzerlik gösteriyorsa geçerli sonuçlar türetilir. Kiaer eğer herhangi bir kısmi araştırma kontrol edilebilen faktörlerde doğru ise büyük olasılıkla kontrol edilemeyen faktörlerde de doğru olduğunu ve farklı metotlarla aynı sonuçlar elde edilebiliyorsa sonuçlara büyük güven duyulabileceğini varsaymıştır (Seng, 1951).

Bazı istatistikçiler ılımlı baksa bile Kiaer'in vizyonu kendi zamanının çok ötesinde olduğundan dolayı önerisine ilk tepki oldukça negatif olmuştur. Ana muhalefet Münih Üniversitesinden etkili bir Bavyeralı istatistikçi olan Georg von Mayr'den (1841–1925) gelmiştir. İtalyan İstatistikçi Luigi Bodio (1840–1920) da Mayr'ın görüşünü

desteklemiş ve Avusturyalı istatistikçi Heinrich Rauchberg de konu ile ilgili daha fazla tartışmanın gereksiz olduğunu belirtmişlerdir (Bethlehem, 2009).

Kiaer 1895’de öne sürdüğü fikirlerini ISI’nın 1897’de St. Petersburg’daki toplantısında da yinelemiştir ve biraz daha açmıştır. Temsil edici araştırma ile çok sayıda yerleşime yayılan kısmi araştırmaları anladığını belirtmiştir. Bunlar tüm alan üzerine yayılmıştır böylece tümün bir minyatürünü düzenlerler. Yerleşimler keyfi olarak yapılmamıştır, bunun yerine tam sayım sonuçlarına dayanarak oransal bir gruplama yapılmıştır. Sonuçlar bu sayımlarla karşılaştırılarak kontrol edilmelidir (Kruskal ve Mosteller, 1980).

Kiaer 1897’de İskandinavyalı istatistikçilerin Stockholm konferansına katılmıştır. Bu konferans temsil edici metotların özellikle sosyal istatistiklerde temsil edici olduğu konusunda görüş birliğine varmış fakat mümkün olduğunda tam sayıya dayalı istatistiksel araştırmaların devam etmesi konusunda ısrar etmiştir. Ayrıca bir temsil edici kısmi araştırma uygulandığında bir örneğin nasıl seçildiği ile ilgili bir tam rapor yayınlanması konusunda karara varılmıştır (Kruskal ve Mosteller, 1980).

Seng’de belirtildiği gibi, Kiaer yine Allgemeines Statistisches Archiv (1899) de bir makalesinde toplantıda ifade ettiği durumları yinelemiştir. Bu makalenin göze çarpan noktaları aşağıdaki biçimde özetlenebilir (Seng, 1951):

- 1) Temsil edici metot sadece sosyal ve ekonomik araştırmalar alanında değil aynı zamanda tarımsal alanda ve ormancılık alanında da uygulanabilir.
- 2) Tam bir temsil edici seçim elde etmek için araştırma altındaki farklı toplulukları gruplamak gereklidir. Böylece sosyal incelemelerde kasabalar ve kırsal bölgedeki topluluklar birbirinden ayırt edilebilir ve ayrıca kendi içlerinde büyük, orta ve küçük, kıyı şeridi, iç kısım, endüstriyel ve kırsal alan olarak da ayrılabilir. Gerçekten temsil edici bir örnek için bir ülkeyi homojen kısımlara gruplama (veya tabakalama) işlemi dikkatlice uygulanmalıdır.
- 3) Eğer mümkünse bir temsil edici örnek elde etmede iki veya daha fazla farklı sistem kullanılmalıdır, böylece araştırmanın sonuçlarına daha fazla güvenilebilir ve metodun faydasının ispatı elde edilebilir.
- 4) Metodun pratik ve teorik yönlerini çalışmak ve geliştirmek önemlidir, böylece temsil edici istatistiklere uygun limitler yerleştirilebilir.

Örnekleme dayalı araştırmalarla ilgili yapılan daha ileri çalışmalara değinilmeden önce dikkat edilmesi gereken bir konu vardır. Araştırmalarda rastsallaştırmanın (randomization) rolü ile ilgili çok az veya hiçbir tartışma yapılmamıştır. Bellhouse (1988)’de belirtildiği gibi, Kiaer (1897) rastsallığın rolünü şu kelimelerle ele almıştır “partiler (lots) halinde seçilmiş örnek”. Fakat bu fikri herhangi bir makalesinde geliştirmemiştir.

Kiaer 1901’de Budapeşte’de yapılan ISI toplantısında yeni vurgularla belirttiği konulara geri dönmüştür. Burada Kiaer ilk olarak coğrafi birimlerin ülke geneline yayılmasına olan ihtiyacı vurgulamıştır. Ayrıca eğer birimler gelişigüzel bir biçimde toplanırsa birçoğunun bir bölgeye veya birkaçına düşebileceğini ve böyle bir kısmi araştırmanın temsil edici olamayacağını belirtmiştir. Burada gelişigüzel (au hasard) kelimesi herhangi bir olasılıksal ima taşımamaktadır (Kruskal ve Mosteller, 1980).

Herhangi bir ilgili tamsayım olmadan karşılaştırma yaparak temsil ediciliği kontrol etmek için, Kiaer araştırmayı her biri kendi içinde temsil edici araştırma olan iki veya üç ayrı kısma bölmeyi önermiştir. Eğer bu kısımlar benzer sonuçlar verirse, böylece kazanılan geçerlilik inkâr edilemez. Bu kısımların Kiaer'in standartlarına göre temsil edici olduğu tam olarak bilinemese de bugün alt örnekleme (subsampling) ve tekrar (replication) olarak adlandırılan kavramların önceki bir notasyonu görülmektedir (Kruskal ve Mosteller, 1980).

Kiaer'in konuşmasından sonraki tartışmada Ladislaus Josephowitsch Bortkiewicz (1868–1931) önemli bir katkıda bulunmuştur. Bortkiewicz “temsil edicilik” terimini hak etmek için örnek ve kitle arasında yaklaşık bir uyum gereksinimi sağlamada bir önem (significance) testi önermiştir ve şöyle belirtmiştir “Gözlenen farklar rastsal olarak ele alınabilir mi?” (Kruskal ve Mosteller, 1980).

Benzer durumlar için ihtiyaç duyulan formüllere Poisson ile sonuç çıkarılabilir mi; bu rastlantı sonucu olan sebeplerle erişilebilecek iki oran veya katsayı arasında ne kadar fark olduğunun kolayca hesaplanmasına imkân verir. Bortkiewicz'e göre Kiaer'in durumunda kitle ve örnek arasında gözlenen farklar bu limitlerin dışında bulunuyorsa, o zaman temsil edicilik başarısızlığa uğrar. Diğer durumda sağlanır. Bu yaklaşımı Kiaer'in “Allgemeines Statistisches Archiv” makalesindeki örneklerle uyguladığında gözlenen farkların kabul edilebilir limitlerin oldukça dışında olduğunu görmüştür. Böylece örnekleri seçmedeki Kiaer'in büyük yeteneğinin teorik olarak gereksinim duyulan doğrulukla sonuçlanmadığını belirtmiştir. Bunun sonucu olarak bu örnekler için “temsil edici metot” terimini reddetmiştir Boertkiewicz'in aktarması öncelikle temsil edici metotla ilgili ISI makalelerine olasılığın ilk girişidir. Kiaer bazı küçük itirazlara cevap vermiş ancak hiçbir yerde istatistiksel önem konusunda Boertkiewicz'e verdiği bir cevaba rastlanmamıştır (Kruskal ve Mosteller, 1980).

1903 Berlin ISI toplantısının raporu Kiaer'in önceki tekrarlanan ifadelerine ek olarak şöyle bir yenilikçi ifadeye de sahiptir: “Temsil edici metodun sonuçlarının yalnızca yaklaşımlar için olduğu doğrudur, ancak bu tüm istatistikler için de geçerlidir. Buna tam sayımdan çıkarılan sonuçlar da dâhildir. Bir temsil edici örnek ile kitledeki oranlar arasındaki farklar bir tam sayımdan diğerine olan farklardan çok az bir farkla büyüktür.” Bu kitlenin kendisinin zamanla değişkenliğe uğradığı ile ilgili ilk açık ifadedir (Kruskal ve Mosteller, 1980).

ISI toplantısındaki konuşmacılar temsil edici örneklemenin benzer kritiklerini ve savunmalarını gözden geçirmişlerdir. Burada özellikle Lucien March'ın (1859–1933) katkısı önemlidir. İlk olarak March temsil edici metodun uzun zamandır kullanıldığını belirtmiştir ve 1903'de yapılan toplantıda rastsallaştırmanın kısmi araştırmalara objektif bir baz sağlayabileceğini belirtmiştir (Pfeffermann ve Rao, 2009).

March rastsallığı belirtmek üzere ‘au hasard’ kelimesini kullanmıştır. March kimsenin temsil edici metodu sağlamak için rastsallığa olan ihtiyacı belirtmediğini söylemiştir. March'a göre rastsal bir örneği ele almanın (‘prendre au hasard’- rastsal olarak alınması) farklı anlamları olabilir. Örneğin sonlu bir kitleden iadesiz yapılan örnekleme kastedilebilir veya kümelerin iadesiz olarak basit rastsal örnekleme ile seçildiği küme örneklemesi kastedilebilir (Kruskal ve Mosteller, 1980; Bellhouse, 1988). March yorumlarında tabakalamayı önerme şansını kaçırmıştır. (Kruskal ve Mosteller, 1980).

Böylece olasılıksal örneklemeyle ilişkin fikirler rastsallaştırma ile birlikte derece derece girmeye başlamıştır, ancak daha çok açık ve inandırıcı değildir. March sadece rastsallaştırma yaklaşımları ile ilgili katkılarından dolayı değil ayrıca 1903 Berlin ve 1925 Roma ISI toplantılarını düzenleyenlerden biri olduğu için de önemlidir.

Buraya kadar olan kısmı özetlemek gerekirse, Kiaer 1895, 1897, 1903’ de yapılan tüm toplantılarda örneklemeyle savunmuştur. Bazı kuvvetli destekler ve bazı kuvvetli muhaliflerle karşılaşmıştır. Kiaer’in savunduğu örnekleme günümüz modern örnekleme teorisi metotlarına benzerdir. Ancak aralarındaki temel fark açık bir rastsal seçim mekanizması yoktur. Yapılan çalışmalarda daha çok temsil edicilik kısıtı altında amaca yönelik örnekleme (purposively) kullanılmaktadır. Boertkiewicz temsil edicilik için bir önem testi önermiş ve uygulamıştır. Fakat Kiaer bir olasılıksal modele sahip olmadığından bu test spesifik olarak kalmıştır. March tartışmaları genişletmiş ve olasılıksal örnekleme fikrine girişi genişletmiştir.

1903–1925 yılları arasında ISI Bülteninin sayfalarında örneklemeyle dayalı araştırmalarla ilgili çok az şeye rastlanmaktadır.

Ancak bu dönemde örneklemeyle dayalı araştırmalara en büyük desteği verenlerden biri Arthur Lyon Bowley (1869–1957) olmuştur. Bowley 1906’da yayınladığı bir makalesinde şöyle belirtmiştir (Bowley, 1906): “Örnekleme metodu tabii ki istatistiğe olasılık teorisini uygulamanın örneklerinden sadece biridir. Bunu ayrıntılı olarak ele aldım, çünkü bu metot ısrarla göz ardı edilmiştir ve hatta kullanıldığında da duyarlılık testi göz ardı edilmiştir. Böylece araştırmaların çok güçlü bir silahı bir tarafa atılmıştır. Sıklıkla sayımlarda yapıldığı gibi tüm bir alanı kapsamak mümkün değildir ve gerekli de değildir. Örneklemeyle kullanarak da iyi sonuçlar elde edilebilir ve çok sıklıkla küçük örnekler yeterlidir. Tek güçlük araştırmada her bir bireyin ya da birimin eşit seçilme şansına sahip olmasını garanti etmektir”.

Bowley 1906’daki çalışmasında örneklemeyle uygulanabilirliğini şöyle bir deneyle açıklamıştır. 3878 şirketin karlarının bir listesi olan Investor’s Record’dan 400 hacimli bir örnek seçilmiştir. Seçim Nautical Almanac’dan bir tablodaki son haneleri kullanarak olmuştur. Bunlar rastsal sayılar olarak değerlendirilmiştir. Eğer sayı 3878’den büyük ise göz ardı edilir, eğer altında ise (örneğin 0063 gibi) örnekleme yapılan tablodan ilgilenilen değer yazılır. Böylece 3878 girişin her birine eşit seçilme şansı verilir. Bu şekilde örneklenecek grup birimlerinin her birinin eşit seçilme olasılığına sahip olmasını ve ilgilenilen değişkenin aldığı değerden bağımsız olarak ele alınmasını garanti etmek önemlidir (Aldrich John, 2008). Burada 400 örneklenen birim önce kayıt edilmiş ve daha sonra peş peşe 10 hacimli 40 gruba ayrılmıştır. Amacı basit rastsal örnekleme için merkezi limit teoremini doğrulamaktır. 40 gruba ait örnek ortalamalarının dağılımı normal eğri ile karşılaştırılır ve yeterli derecede normal eğri tarafından tanımlandığı görülür. Bowley elde ettiği sonuçların örnekleme yaptığı kitle hacminden bağımsız olduğunu belirtmiştir. Bu sonuç tam olarak doğru değildir, çünkü örnek ortalamasının varyansında sonlu kitle düzeltme faktörünü dikkate almamıştır. Bellhouse’da belirtildiği gibi düzeltme faktörünü de dikkate alan sonuç birkaç yıl sonra Isserlis (1918) tarafından elde edilmiştir (Bellhouse, 1988). Bowley örnekleme için basit rastsal örneklemeyle matematiğini kullanmıştır ve örneklemeyle olasılıksal hatalarının ortalama etrafındaki artı ve eksi üç katı aralığındaki tahminlerini önermiştir. Bu yaklaşık olarak iki standart hatadır ve bununla birlikte tam olarak olmasa da %95 güven aralıklarını sağlar (Kruskal ve Mosteller, 1980).

1909'da Bowley Elementer İstatistik Elkitabında Örneklemeye ile ilgili özet bir bölüm koymuştur. Ticarete, madencilikte ve endüstride örneklemeye genel referanslar yapmıştır. Bowley'in önerileri uygulandığında Britanya'da çalışılan yoksulluk ile ilgili yolları değiştirmiştir. 1912'de yoksullukla ilgili bir araştırmada her bir yirminci çalışan sınıfını almıştır ve 1913'de bunu 3 şehir için daha genelleştirmiştir. Örneklemenin olasılıksal hatalarını hesaplamış ve görüşmelerde, tanımlamalarda ve tahminleme süreçlerinde diğer yapılan hata kaynaklarının farkına varmıştır. Böylece istatistiksel teori ve araştırma tasarımını birlikte ele almıştır (Stephan, 1948; Biemer ve Lyberg, 2003).

Bowley bir araştırmada sosyal koşullarla ilgili olarak dört hata kaynağını ele almıştır (Bowley ve Hurst, 1915):

- 1-Toplanan bilgi her zaman tam olarak doğru olmayabilir,
- 2- Kullanılan tanımlar ve standartlar gevşek, uygunsuz veya yanlış tasarlanmış olabilir,
- 3-Aktif olarak ziyaret edilen haneler kitlenin uygun (fair) bir örneğini içermeyebilir,
- 4-Bir kısmı ölçerek tümü tahminleme sürecinden kaynaklanan hesaplanabilir olası hatalar da vardır.

John Hilton "Some Further Enquires by Sample" isimli makalesinde önceden yapılmış olan dört incelemeyle ilgili istatistiksel problemleri, deneyimleri ve çıkarılan dersleri ele almıştır ve yukarıda belirlenen sapma kaynaklarının gerçekte ortaya çıktığını raporlamıştır (Hilton, 1928).

Bowley birçok makalede rastsal örneklemeyi ve başlıca örneklemeye çerçevelerine duyulan ihtiyacı ele almıştır. Amaca yönelik seçim için bir teoremin ana hatlarını oluşturmuş ve araştırma tasarımı için kılavuz noktalar sağlanmıştır. Ancak dikkat edilmelidir ki ne Kier ne de Bowley tüm aşamalarda tam olarak rastsallaştırmayı savunmamışlardır. Onlar öncelikle rastsal ve amaca yönelik seçimin bir karışımını savunmuşlardır. Örneğin amaca yönelik büyük kümelerin seçilmesi durumunda rastsal veya gelişigüzel olarak birimlerin ve küçük kümelerin seçilmesi gerekmektedir (Biemer ve Lyberg, 2003).

3.3 1924–1934 Dönemi: ISI Komisyonu

1924 yılında ISI istatistikte temsil edici metot uygulamalarının çalışılması amacı için bir komisyon atamıştır. Danimarkalı istatistikçi Adolph Jensen bu komisyonun raportörü olarak atanmıştır ve diğer üyeler Arthur Lyon Bowley, Corrada Gini, Lucien March, Verrijin Stuart ve Franz Zizek'dir.

Komisyon bulgularını bir diğer ifadeyle 1925'de Roma'daki ISI toplantısında alınan kararları 1926'da raporlamıştır. Burada son birkaç itiraz edenin gönlünü de almak için dikkatlice bir üslup kullanılarak şöyle bir karar benimsenmiştir: "Hem rastsal (kitlenin tüm elemanları eşit olasılıkla seçilebilmelidir) ve hem de amaca yönelik seçimle belirlenmiş örneklemeye metotları (küme örneklemesi gibi, fakat görünürde daha büyük kümelerle, seçilen kümelerin tam listesi ve bu seçimi yapmada yargı) kabul görecektir." (Kruskal ve Mosteller, 1980; Bellhouse, 1988).

Rastsal örneklemenin tanımları genel olarak tamamlanmamıştır. Her bir birimin örnekte görünme olasılığının eşit olmasına gereksinim duyar, fakat ortak (joint) olasılık

konusunda bir fikir belirtmemişlerdir. Tüm örnek noktalarının dağılımına bakılması fikri halen geleceğe uzanmaktadır. March'ın önceki katkılarının dışında olasılıksal örneklemenin genel fikri henüz doğmamış görünmektedir. Tabakalı örnekleme ile ilgili bazı tartışmalar olmasına rağmen, daima tabaka örnek hacimleri birimler üzerinde sabit seçim olasılığını yönetmek için tabaka hacimlerine orantısaldır (Kruskal ve Mosteller, 1980).

Komisyon raporu için ana konuşmacılar Jensen ve Bowley'dir. Jensen pratikte temsil edici metodun uzun bir tanımını sağlamıştır.

Kruskal ve Mosteller(1980)'da belirttiği gibi Jensen (1926a) tarafından yazılan iki makale geniş tartışmalar sağlar. Bunlardan ilki ISI komisyonunun raporudur. Amaca yönelik örnekleme fikrini biraz daha ayrıntılı olarak açıklamıştır. Jensen'e göre seçilen gruplar kitlenin hâlihazırda bilinen karakteristikleri için (kontrol karakteristikleri) kitlenin ortalamalarıyla yaklaşık olarak eşit ortalamalara sahip olmalıdır. Ayrıca Jensen seçilen grupların kontrol değişkenlerinin, sadece ortalamalarının değil, değişkenlerin dağılımlarının da kitle ile benzer dağılımlara da sahip olmasına gereksinim duymuştur. İkinci makalesinde (1926b), Jensen pratikte temsil edici metodu mantıklı bir uzunlukta açıklamış ve uzun bir bibliyografi vermiştir. Bu makale küme örneklemesinin nispeten açık bir tartışmasını içermektedir (grupların rastsal seçimi), ancak uygulamada rastsal seçim daha çok sistematik şekilde yapılmış gibi görünmektedir. Büyük grupların amaca yönelik seçimini ve sonra gruplama içerisinde birimlerin rastsal seçimini izlemektedir (Kruskal ve Mosteller, 1980).

Bowley'in belki de bilinen en önemli katkıları 1924'de oluşturulan komisyon aracılığıyla olmuştur ve Bowley de rastsal ve amaca yönelik seçimde bilinen sonuçları özetleyen bir teorik monografi sağlamıştır. Birkaç diğer fikre ilaveten bu monografi orantısal tahsislemeli tabakalı örneklemenin gelişimini (bazen Bowley allocation olarak da bahsedilir) ve ilgilenilen değişken ve kontrol değişkenleri arasındaki korelasyonlar aracılığıyla amaca yönelik seçimin teorik gelişimini içerir. İkinci olarak da bir amaca yönelik örnekleme tasarımı altında tahminin duyarlılığının ölçümü için formülü içerir. Orantısal tahsisleme konusundaki çalışması Bowley'in rastsallaştırılmış örnekleme tasarımında (randomized sampling design) tüm birimler için eşit seçilme olasılıklarını yönetme konusunda çalışmaya devam etme arzusunu yansıtmıştır (Bellhouse, 1988). Bowley ayrıca yukarıdaki gelişimlere ek olarak birkaç yönden de yeni çığır açmıştır. Neredeyse güven aralığı fikrine ulaşmıştır ve daha da ileri giderek neredeyse çoklu karşılaştırma fikrine ulaşmıştır- yani güven bölgesi fikrine ulaşmıştır- ve daha sonra bir Bayesian muameleyi ortaya koymuştur (Kruskal ve Mosteller, 1980).

Hollanda İstatistik Müdürü olan Verrijn Stuart birkaç dikkate değer yorumda bulunmuştur. Bir örnekleme dayalı araştırmanın hiçbir zaman bir tamsayıdan daha iyi olamayacağını belirtmiş ancak bir tamsayımı uygulamanın çok güç olduğunu ve bu durumda bir örnekleme dayalı araştırmanın kullanışlı olabileceğini belirtmiştir (Bethlehem, 2009).

Bazı eleştirisel açıklamalarına rağmen, Verrijn Stuart sıklıkla kaliteyi kaybetmeden örnekleme dayalı araştırmaları kullanmanın mümkün olduğuna karar vermiştir. Her bir araştırma için genel bir öneri verememiştir. Mümkün uygulamalar daima dikkatli bir biçimde test edilmelidir. Detaylı dokümantasyonlar üretilmelidir. Böylece bir süre sonra daima kesin sonuçların nasıl elde edilebileceği öğrenilebilir.

Örnekleme konusunda Verrijn Stuart rastsal seçimi savunmaktadır. Ona göre amaca yönelik örnekleme daima sübjektif kararlar yansıtır. Bundan rastsal örnekler kullanılarak kaçınılabılır (Bethlehem, 2009).

1925'deki toplantıda tartışma artık örnekleme yapıp yapmamak değil, bunun nasıl yapılacağı olmuştur. Hem tabakalı da ve hem de küme örneklemesinde ilerlemeler görülmüştür. Belirsizlikler kalmasına rağmen, amaca yönelik örnekleme fikri yaşamaya devam etmiş ve uygun bulunmuştur. Bu komisyonun olumlu görüşlerine rağmen yinede örneklemenin kabul görmesi bazı ülkelerde onlarca yıl almıştır (Kruskal ve Mosteller, 1980).

Bu çalışmalardan bağımsız olarak Rusya'da aynı dönemlerde benzer çalışmalar yapılmaktadır. Rusya'nın tarihindeki örneklemeyle dayalı çalışmalarla ilgilenen iki Chuprov vardır. Biri Aleksander Aleksandrovich Chuprov'un babası olan Aleksander Ivanovich Chuprov (1842–1908)'dur. A.I. Chuprov Moskova Üniversitesinde politik ekonomi ve istatistik profesörüdür ve Zemstvo istatistiklerinin arkasındaki itici güçtür.

Aleksander Aleksandrovich Chuprov (Tschuprow) (1874–1926) da St. Petersburg Polytechnic Enstitüsünde Ekonomi bölümünde Profesör olarak çalışmıştır. Devrim öncesi Rusya'sında istatistiğe itici bir güç olmuştur. Örneklemeyle dayalı araştırmalar alanında yaptıklarından bazıları hem SSCB içinde ve hem de dışında seçkin olmuşlardır. Tschuprow örneklemeyle dayalı araştırmalar ve tabakalı rastsal örnekleme altında tahminler için formüller geliştirmiştir. Neyman'ın sonuçlarını önceden sezinlemiştir. (Heyde et al. 2001; Biemer ve Lyberg, 2003). Rusya'da bu dönemlerde çalışan daha birçok ehil istatistikçi olmuştur. Bunlar arasında A.Kaufman, V.Romanovsky, E.Slutsky, B.Iastremsky, Boiarsky, Tchetwerikov, Obukhov, Kowalsky v.b. sayılabilir. Bunların birçoğu gençtir ve sonradan meydana gelen olaylar onların Rus sınırları dışında tanınmasını engellemiştir. Rus istatistikçilerin birçoğu da, Tchetwerikov gibi, örnekleme birimlerinin rastsal olarak seçilmesi gerektiğine inanmışlar ve örnek hacmindense örnek birimlerinin seçilme yolunun önemli olduğunu belirtmişlerdir (Zarkovich, 1956; 1962).

3.4 1934 ve Sonrası: Neyman Dönüm Noktası ve Sonraki Gelişmeler

Amaca yönelik örnekleme ve rastsal seçim metodunun birlikte kullanılması 1934'e kadar sürmüştür. 1934 yılında Polonyalı bilim adamı Jerzy Neyman (1894–1981) örneklemeyle dayalı araştırmalarda bir dönüm noktası olan "On Two Different Aspects of the Representative Method: The Method of Stratified Sampling and the Method of Purposive Selection" isimli makalesini Londra'daki Kraliyet İstatistik Topluluğu (Royal Statistical Society) toplantısı için sunmuştur. Bu makale istatistik sahasına önemli bir katkı olarak yerini almıştır. Bu makale olasılıksal örnekleme olarak bilinen kavramın teorik araştırmalarında, metodolojik gelişimlerinde ve uygulamalarında teşvik edici olmuştur. Bu makalede Neyman rastsal örneklemenin önemini vurgulamıştır ve güven aralıkları kavramına dayanan yeni bir tahmin teorisi geliştirmiştir. Optimum tabakalama, küme örnekleme, büyük örneklerde doğrusal tahmincilere normallik yaklaşımı ve amaca yönelik seçim modeli ile ilgilenir. (Neyman, 1934; Hansen, 1987; Bellhouse, 1988; Beaud ve Prevost, 1998; Biemer ve Lyberg, 2003; Bethlehem, 2009). Neyman rastsallaştırma aracılığıyla kitle hakkında yorumlamalar yapmanın mümkün olduğunu belirtmektedir. Ayrıca amaca yönelik örneklemede yorumlamalara karşılık gelen bir istatistiksel teorinin bulunmadığını belirtmiştir. Örneğin rastsal örneklemeyle

kullanarak istatistiksel güven aralıklarını oluşturmak mümkün iken amaca yönelik örnekleme için bu mümkün değildir (Biemer ve Lyberg, 2003).

Neyman'ın makalesinin önemli bir katkısı da örnekleme prosedürünün tam belirlenmesinde olan ısrarıdır. Önceden amaca yönelik örnekleme için kullanılan prosedürler ve olasılığın girip girmediği veya nasıl girdiği açık olarak belli değildir. Neyman kümelerin amaca yönelik örnekleme probleminde etkileyici bir çözüm getirmiş olsa bile, amaca yönelik örnekleme cazibesini yitirmeye başlamıştır (Kruskal ve Mosteller, 1980).

Hem Bowley önceki çalışmalarında amaca yönelik seçim hakkında eleştiride bulunmuş ve hem de Tschuprov daha önce tabakalı örneklemede optimum tahsisleme ile ilgili bazı sonuçlar sunmuş olduğundan dolayı bu makale özellikle ilgi çekmiştir.

Tschuprov tarafından elde edilen sonuçların bazıları Neyman tarafından geliştirilmiş görünmektedir. Neyman sonlu kitleden örnekleme ile ilgili bir teori geliştirdiğinde Tschuprov'un sonuçlarına erişip erişmediği açık değildir. Sonuçlar bazı alanlarda üst üste çakışmaktadır. Ancak ilgi çekici olan şudur ki Neyman 1920'lerden önceki çalışmalarını sunduğunda hiçbir zaman Rus bilim adamından bahsetmemiştir (Biemer ve Lyberg, 2003).

Tschuprov'un çalışmasının Neyman'ın çalışması kadar bir patlama yapamamasının temel iki nedeni olduğu söylenebilir. Birincisi Tschuprov, Neyman'ın sağladığı cevabın sadece bir kısmını vermiştir. Sabit içerilme olasılıkları kalıbını kırmıştır ancak metodun ana rakibi olan amaca yönelik seçime göre üstünlüğünü gösterememiştir. Tschuprov'un etkisinin çok olamamasının diğer bir nedeni de kendi stilindeki formüllerinden, özetlemesinden ve coğrafi hareketliliğinden kaynaklanmaktadır. Tschuprov'un Rus devrimini izleyerek coğrafi hareketliliği çok olmuştur. Mayıs 1917'den sonra hiç Rusya'da bulunmamıştır. Oysaki Neyman daha kolay ve anlaşılır bir dil kullanmıştır ve Tschuprov gibi hareketli bir yaşam tarzına sahip değildir. Ancak bildiklerini anlatmak ve eğitimler vermek üzere çeşitli seyahatler yapmıştır (Bellhouse, 1988). Bu seyahatler sırasında Washington'da yapılan "İnsan Kitesinin Örnekleme" (Sampling Human Populations) ile ilgili bir konferansta Mr. Milton Friedman ve Dr. Sidney Wilcox Neyman'a çifte örnekleme (double sampling) ile ilgili soru sormuşlardır. Neyman bu soruya 1938'de yazdığı "Contribution to the Theory of Sampling Human Population" isimli makalesi ile yanıt vermiş ve ayrıca bu makale ile örnekleme dayalı araştırmalar teorisine maliyet fonksiyonlarının kullanımını sokmuştur (Neyman, 1938).

Neyman'ın çalışmalarının orijini tarımsal istatistiklere dayanmaktadır ve bu Rothamsted'daki Ronald Aymles Fisher (1890–1962) tarafından yapılan deney tasarımı çalışmaları için geçerlidir. Fisher'in çalışmalarının ve rastsal deneylerdeki fikirlerinin örnekleme dayalı araştırmalar için önemi büyüktür. Fisher'in tarımsal deneylerde uyguladıkları özellikle rastsallaştırma ve içerilme (inclusion) olasılıkları hakkında örnekleme dayalı sosyal araştırmalara çok yön gösterici olmuştur. Ancak tüm zamanlarda istatistik teorisine büyük katkı sağlayan iki kişi olan Neyman ve Fisher arasındaki düşmanlık örnekleme dayalı araştırmaların gelişimini belki de önemli ölçüde zayıflatmıştır (Biemer ve Lyberg, 2003). Fisher hiçbir zaman makalelerinde 'temsil edici örnek' terimini kullanmamıştır. Ancak makalelerinde şöyle alıntılara rastlanmaktadır. Fisher (1936): "Canlılar üzerinde çalışan kişilerin denekleri seçme ve yeterince bulabilme şansı vardır. Ancak bir toplum ırksal olarak çok iyi temsil

ediciliği olmayan bir mezarlığında araştırma yapan bir kişinin böyle bir şansı yoktur.” Bu kısıt yeterli örnek hacmine olan ihtiyacı hissettirir.

Kiaer, Bowley ve Neyman’ın başlattığı örneklemeyle dayalı araştırmalar konusunda 1930’lardan itibaren dünyanın başka yerlerinde de örneklemeyle dayalı araştırmalar yapılmaya başlanmıştır. Bu konuda çalışanlardan bir tanesi de Prasanta Chandra Mahalanobis (1893–1972)’dir. Mahalanobis’in en dikkate değer ve en kalıcı katkıları büyük ölçekli örneklemeyle dayalı araştırmaları oluşturmak olmuştur. İstatistiksel teorisinin çeşitli somut Hint problemlerine uygulamasını yapmıştır. Mahalanobis olay yeri istatistiklerinde başarıya ulaşmadan önce neredeyse Hindistan’da bilinmemektedir. Bu konu daha önce Hindistan üniversitelerinde ele alınmamıştır (Gupta, 2009).

Mahalanobis örneklemeyle dayalı araştırmaların öncülerinden biridir. Mahalanobis önce 1920’lerde İstatistiksel Laboratuvarını kurmuş ve daha sonra bunu geliştirerek 17 Aralık 1931’de Hindistan İstatistik Enstitüsünü (the Indian Statistical Institute) kurmuştur. Bu kurum Hindistan’daki sosyal ve ekonomik planlamaya destek sağlamak amacıyla örneklemeyle dayalı araştırmaların gelişimine ve uygulamasına katkıda bulunmuştur. Mahalanobis şu görüşü savunmaktadır: “Gelişmiş ülkelerdeki prosedürleri yalnızca taklit etmemek gereklidir, aynı zamanda Hindistan’a özel koşullara göre de metotlar ayarlanmalıdır. Bu örneklemede uygun yaklaşımdır. Kullanılan metotlar, hizmet verilebilecek belirli amaçlara odaklanmaya ilaveten, ilgili mevcut kaynaklar kullanılmalı ve tanımlanmalıdır. Mahalanobis “istatistiksel mühendislik” felsefesini geliştirmiştir ve birkaç yıl deneysel aşamalarda tasarımlar uygulamıştır, maliyet ve varyans tahminlerine göre tasarımlar geliştirmiş ve kurmuştur. Aynı zamanda toplam araştırma hatalarını (total survey error) kontrol edebilmek için metotlar geliştirmiştir (Hansen, 1987). Uygulamasını tarımsal veri toplayan saha çalışanları ile yapmıştır. Bu hataları yorumlama (interpretation) olarak adlandırılan bir metotla tahminlemiştir. Bu metot günümüzde görüşmecilerden, kodlamalardan ve kodlamalardan kaynaklanan hataları tahmin etmede kullanılmaktadır (Biemer ve Lyberg, 2003). Mahalanobis ayrıca 1950’de Ulusal Örneklemeyle Dayalı Araştırmaları (National Sample Survey) ve Merkezi İstatistik Örgütünü (Central Statistical Organization) (1951) kurmuştur. Mahalanobis 1971’de ölümüne kadar kuvvetini ve fikirlerini yaymadaki cazibesini sürdürmüştür.

Pandurang Vasudeo Sukhatme (1911–1997) de tarımsal örnekleme ile ilgili geniş çaplı araştırmalar yapmıştır. Çeşitli teorik ve pratik katkılarda bulunmuştur, bunların bazıları tabakalı örnekleme üzerinedir. Ayrıca pilot çalışmaların kullanımı, çok aşamalı örnekleme, çifte örneklemede regresyon tahmincileri ve örnekleme dışı hataların kontrolü ile ilgili de önemli katkılar yapmıştır (Hansen, 1987; Som, 1998).

Sukhatme ve Mahalanobis’in pratik istatistiksel uygulamalarının bazılarında güçlü derecede anlaşmazlıkları vardır ve bunları bazı tartışmalara taşımışlardır. Daha sonra, Sukhatme Roma’daki Besin ve Tarımsal organizasyona (Food and Agriculture Organization) gitmiştir ve burada araştırmalardaki ölçüm hataları ve örnekleme ile ilgili teori ve metotlarını geliştirmeye devam etmiştir (Hansen, 1987).

1942 makalesinde Jessen iki ardışık durumda kitle karakteristiklerinin tahmin edilmesi için örneklemeyle dayalı araştırmaların tasarımı ile ilgilenmiştir ve araştırma tasarımını detaylı olarak açıklamıştır. Ayrıca örneklemeyle dayalı araştırmalarda yapılan hata kaynaklarına değinmiş ve örnekleme hatalarını hesaplamıştır. Burada iki durumun her

birinde bir örneğin eşleştirilmiş ve eşleştirilmemiş kısımlarının optimum tahsisi için bir teori sunmuştur. Bu öncül adım sonradan zaman serisi tahmini için ardışık durumlarda alınan araştırmalar için örneklerin döndürülmesinde (rotate) teori ve dizaynların gelişimi için uygulanmıştır (Jessen, 1942; Hansen, 1987).

Ayrıca yine Hansen’de belirtildiği gibi 1942’de Cochran örnekleme dayalı araştırmalarda regresyon tahmininin kullanımına özel bir katkı yapmıştır ve bu çalışma daha sonra Madow ve Madow tarafından geliştirilmiştir (Hansen, 1987).

1943’de Hansen ve Hurwitz mevcut kaynakların en etkili biçimde kullanılması amacıyla geliştirilmiş olan küme örnekleme, altörnekleme ve çifte örnekleme gibi metotlara alternatif olarak kendi çok aşamalı örnekleme teorilerini önermişlerdir. Bu tasarım bir tabakalı kitlenin elemanlarının kümelerinin örnekleme ve seçilen kümelerin her birinden elemanların alt örnekleme içerir. Bir tabaka içerisindeki öncül (primary) örnekleme birimlerinin her birindeki eleman sayıları eşittir (Hansen ve Hurwitz, 1943).

Bethlehem’e göre örnekleme dayalı araştırmaların klasik teorisi az çok 1952’de Horvitz ve Thompson ile tamamlanmıştır. Horvitz ve Thompson sapmasız tahminler oluşturmak için genel bir teori geliştirmişlerdir. Seçim olasılıkları ne olursa olsun bilindikleri ve pozitif oldukları müddetçe, kullanışlı tahminleri oluşturmak da mümkün hale gelir. Horvitz ve Thompson klasik teoriyi tamamlamıştır ve rastsal örnekleme hemen her yerde çoğunlukla kabul edilmiştir (Bethlehem, 2009).

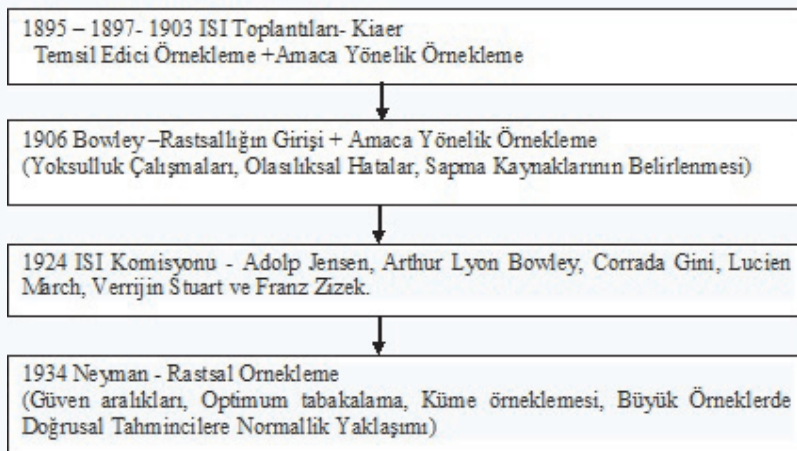
Örnekleme dayalı araştırmaların dönüm noktası Kiaer, Bowley ve Neyman’ın çalışmaları olsa da yukarıda adı geçen diğer birçok bilim adamının da örnekleme teorisinin gelişimine ve uygulamalarına katkıları büyük olmuştur. Özetle, Hansen, Hurwitz, Madow, Cochran, Kish ve Deming modern örnekleme teorisi için baz olan çerçeveyi formüle etmişlerdir. Yates, Sukhatme, Murty ve Des Raj gibi istatistikçiler bu sahayı ileriye götüren kitaplar yazmışlardır. Dalenius, Godambe, Horvitz, Thompson, Jessen, McCarthy, Waksberg, Pritzker, Stephan, Tepping ve Mahalanobis gibi istatistikçiler örnekleme teorisinin gelişimine ve 1940–1965 periyodu sırasında uygulamalarına büyük katkılarda bulunmuşlardır (Biemer ve Lyberg, 2003).

1965 döneminden sonra gerek basit rastsal örnekleme, tabakalı örnekleme, küme örnekleme, çok aşamalı örnekleme, örnekleme hataları gibi konularda ve gerekse örnekleme dayalı sosyal araştırmalar, bu araştırmalarda karşılaşılan hata kaynaklarının incelenmesi ve belirlenmesi ile ilgili konularda çok yoğun çalışmalar ve gelişmeler olmuştur. Türkiye’de de bu konuda çalışmalar devam etmektedir ve örnekleme teorisine ve uygulamalarına katkı sağlayan pek çok önemli çalışma bulunmaktadır. Bunlardan bazıları şu şekilde özetlenebilir: Özellikle basit rastsal ve tabakalı rastsal örneklemede tahmin edicilerin duyarlılığını geliştirmede sıklıkla kullanılan yardımcı (auxiliary) değişkenleri kullanarak kitle ortalaması ve varyansı için geleneksel oransal (ratio) ve regresyon tahmin edicilerine alternatif olarak daha etkin yeni tahmin ediciler geliştirilmiştir (Kadılar ve Çingı, 2003, 2004, 2005, 2006a, 2006b, 2007 ve Koyuncu ve Kadılar, 2009a, 2009b, 2010). Ayrıca kayıp veriler durumunda etkili tahmin edicilerin elde edilmesi gibi konularda da önemli katkılar da sağlanmıştır (Kadılar ve Çingı, 2008). Ayrıca anketler aracılığıyla örnekleme dayalı yapılan araştırmalarda insan faktöründen, veri toplama metotlarından, örnekleme çerçevesinin tam olarak uygun bir şekilde belirlenmemesinden v.b. kaynaklanan birçok sapma

meydana gelmektedir. Bu sapmalar çalışmaların sonuçlarını önemli derecede etkilemektedir. Bu nedenle örnekleme tasarımı ve veri toplama prosedürlerini inceleyen çalışmalar da yoğun bir biçimde sürmektedir. Ayhan (2003) cevaplama hatalarının (response error) modellenmesi ile ilgili ayrıntılı bir çalışmayı içermektedir. Ayhan ve Işıksal (2004), örneklemeyle dayalı araştırmalarda araştırma sorularına cevap vermede insan hafızasının yanıltmasından kaynaklanan hataların etkilerinin tahmin edilmesi için yöntemler ve modeller geliştirmişlerdir. Ayhan (2005)'de internet aracılığıyla veri toplamak için mükemmel bir örnekleme çerçevesi oluşturmada karşılaşılan güçlüklerden bahsedilmektedir. Ayhan (2005)'de belirtildiği gibi Ayhan (2000) ve Ayhan ve Ekni (2003)'de örnekleme çerçevesi (sampling frame) ve kapsam hatası modelleri (coverage error models) geniş bir biçimde ele alınmıştır. Ayhan ve İslam (2005) hane halkı telefon araştırmalarında rastsal numara çevirme (random digit dialing) için geliştirilen örnekleme tasarımlarını ele almışlardır. Örneklemeyle dayalı araştırmalarda karşılaşılabilecek durumlar ve etkileri ile ilgili yapılan çalışmalar hakkında detaylı bir literatüre <http://www.stat.metu.edu.tr/people/faculty/oztas> adresinden de ulaşılabilir. Günümüzde yapılan çalışmalarla ilgili yapılacak olan ayrıntılı bir literatür çalışması bu konuda çalışan araştırmacılara yol gösterici olacaktır.

4. SONUÇ

Bugünkü örnekleme teorisinin temelini oluşturan çalışmalar şaşırtıcıdır ki aslında çok da eski bir geçmişe sahip değildir. Yaklaşık 100 yıl önce yeni çıkan bir kavram olan örnekleme bugün istatistiksel olarak güvenilir bir biçimde kitle parametreleri ile ilgili yorumlamalarda kullanılmaktadır. Bu çalışmanın amacı özellikle örnekleme kavramının girişinden kabul edilmesine kadar en önemli dönem olan 1895 Kiaer ile başlayıp 1934'de Neyman ile günümüz örnekleme yöntemlerinin temelini oluşturan süreci ve bu döneme damgasını vuran Kiaer, Bowley ve Neyman'ın katkılarını detaylı olarak verebilmektir. İlgili dönemde yeni bir yöntem olan örneklemeyle olan ilk tepkiler, karşılaşılan güçlükler ve adım adım rastsal örneklemeyle giden yollar o dönemde yapılan çalışmalardan örnekler verilerek açıklanmaya çalışılmıştır. Dönüm noktaları kısaca aşağıdaki tarihsel şema ile özetlenmiştir.



Şekil 1. Örneklemeyle dayalı araştırmalar 1895–1934 tarihsel gelişim süreci

Ayrıca bilim adamlarını bu sürece götüren gelişmeler ve bu dönemden sonraki gelişmelere ve katkısı olan diğer bilim adamlarına ve yaptıkları katkılara da değinilmiştir. Nispeten yeni bir kavram olan örnekleme ile ilgili günümüzde de tüm bilim alanlarında olduğu gibi örnekleme teorisini ve uygulama alanlarını geliştirmeye yönelik hem sosyal bilimlerde ve hem de fen bilimlerinde çok değerli çalışmalar yapılmaya devam edilmektedir. Neyman'dan sonraki dönem için de kapsamlı bir literatür taraması yapılması yeni araştırmalar için yol gösterici olacaktır.

5. KAYNAKLAR

Aldrich, J., 2008. Professor A.L.Bowley's Theory of the Representative Method. Based on a Talk Given at the Sample Surveys and Bayesian Statistics Conference.

Ayhan, H. Ö., 2005. Sample Adjustment Weights for Internet Surveys: Restricted Access versus Voluntary Participation in Recent Developments and Applications in Social Research Methodology (Dijkum C., J Blasius & C Durand, eds.). Leverkusen-Opladen, Germany.

Ayhan, H. Ö., Ekni, S., 2003. Coverage Error in Population Censuses: The case of Turkey. *Survey Methodology*, 29, 155-165.

Ayhan, H. Ö., 2000. Estimators of Vital Events in Dual-Record Systems. *Journal of Applied Statistics*, 27, 157-169.

Ayhan, H. Ö., 2003. Models of Response Error Components in Supervised Interview Reinterview Surveys. *Journal of Applied Statistics*, 30, 1047 – 1054.

Ayhan, H. Ö., Islam M. Q., 2005. Sample Design and Allocation for Random Digit Dialling. *Quality & Quantity*, 39, 625 - 641.

Ayhan, H. Ö., Işıksal, S. 2004. Memory Recall Errors in Retrospective Surveys: A Reverse Record Check Study. *Quality & Quantity*, 38, 475–493.

Bacaer, N., 2011. *A Short History of Mathematical Population Dynamics*. Springer, London.

Beaud, J. P., Prévost J. G., 1998. The Politics of Measurable Precision: The Emergence of Sampling Techniques in Canada's Dominion Bureau of Statistics. *Canadian Historical Review*, 79(4), 691–725.

Bellhouse, D. R., 1988. A Brief History of Random Sampling Methods, *Handbook of Statistics*. Elsevier Science Publishers B. V., 6, 1-14.

Bethlehem, J., 2009. The Rise of Survey Sampling. *Statistics Netherlands, Discussion Paper (09015)*, Heerlen.

Biemer, P. P., Lyberg, L. L., 2003. *Introduction to Survey Quality*. John Wiley and Sons Publication, USA.

Bowley, A. L., 1906. Address to the Economic Science and Statistics Section of British Association for the Advancement of Science, York, 1906. *Journal of the Royal Statistical Society*, 69(3), 540-558.

Bowley, A. L., Burnett Hurst A. R., 1915. *Livelihood and Poverty*. G. Bell and Sons Ltd, London.

Brunt, L., 2001. The Advent of the Sample Survey in the Social Science. *The Statistician*, 50(2), 179–189.

Desrosieres A., 2002. Three Studies on the History of Sampling Surveys: Norway, Russia-USSR, United States. *Science in Context*, 15(3), 377–383.

Fisher, R. A., 1936. The Coefficient of Racial Likeness and the Future of Craniometry. *The Journal of the Royal Anthropological Institute of Great Britain and Ireland*, 66, 57-63.

Graunt, J., 1662. *Natural and Political Observations upon the Bills of Mortality*. Martyn, London.

Gupta, A., 2009. *Bright Sparks: Inspiring Indian Scientists from the Past*. Indian National Science Academy.

Hansen M. H., 1987. Some History and Reminiscences on Survey Sampling. *Statistical Science*, 2(2), 180–190.

Hansen, M. H., Hurwitz, W. N., 1943. On the Theory of Sampling from Finite Populations. *Annals of Mathematical Statistics*, 14, 333–362.

Heyde, C. C., Seneta, E., Crepel, P., Fienberg, S. E., Gani, J., 2001. *Statisticians of the Centuries*. Springer Verlag, New York.

Hilton, J., 1928. Some Further Enquires by Sample. *Journal of the Royal Statistical Society*, 91(4), 519–540.

Horvitz, D. G., Thompson, D. J., 1952. A Generalization of Sampling without Replacement from a Finite Universe. *Journal of the American Statistical Association*, 47, 663–685.

Isserlis, L., 1918. On the Value of a Mean as Calculated from a Sample. *Journal of the Royal Statistical Society*. 81, 75–81.

Jensen, A., 1926a. Report on the Representative Method in Statistics. *Bulletin of the International Statistical Institute*. 22 (1), 359–380.

Jensen, A., 1926 b. The Representative Method in Practice. *Bulletin of the International Statistical Institute*. 22 (1), 381–439.

Jessen, R. J., 1942. Statistical Investigation of a Sample Survey for Obtaining Farm Facts. Agricultural Experiment Station Iowa State College of Agriculture and Mechanic Arts, 304, Ames, Iowa.

Kadılar, C., Çıngı, H., 2003. Ratio Estimators in Stratified Random Sampling. Biometrical Journal, 2, 218-225.

Kadılar, C., Çıngı, H., 2004. Ratio Estimators in Simple Random Sampling. Applied Mathematics and Computation, 151, 893 - 902.

Kadılar, C., Çıngı, H., 2005. A New Estimator Using Two Auxiliary Variables. Applied Mathematics and Computation, 162, 901-908.

Kadılar, C., Çıngı, H., 2006a. Improvement in Estimating the Population Mean in Simple Random Sampling, Applied Mathematics Letters, 19, 75 – 79.

Kadılar, C., Çıngı, H., 2006b. Ratio Estimators for The Population Variance In Simple and Stratified Random Sampling, Applied Mathematics and Computation, 173, 1047 – 1059.

Kadılar, C., Çıngı, H., 2007. Improvement in Variance Estimation in Simple Random Sampling, Communications in Statistics, 36, 2075 - 2081.

Kadılar, C., Çıngı, H., 2008. Estimators for the Population Mean in the Case of Missing Data, Communications in Statistics, 37, 2226 - 2236.

Kiaer, A. N., 1897. The Representative Method of Statistical Surveys (1976, English Translation of the Original Norwegian). Central Burea of Statistics of Norway, Oslo.

Kiaer, A. N., 1899. Sur Les Methodes Representatives Ou Typologiques Appliquees A La Statistique.11. 1-180.

Koyuncu, N., Kadılar C., 2009a. Ratio and Product Estimators in Stratified Random Sampling, Journal of Statistical Planning and Inference, 139, 2552-2558.

Koyuncu, N., Kadılar C., 2009b. Family of Estimators of Population Mean Using Two Auxiliary Variables in Stratified Random Sampling, Communications in Statistics-Theory and Methods, 38, 2398–2417.

Koyuncu, N., Kadılar C., 2010. On Improvement in Estimating Population Mean in Stratified Random Sampling, Journal of Applied Statistics, 37, 999-1013.

Kruskal, W., Mosteller, F., 1980. Representative Sampling, IV: The History of the Concept in Statistics, 1895-1939. International Statistical Review, 48(2), 169-195.

Malthus, T. R., 1798. An Essay on the Principle of Population. Printed for J.Johnson in St.Paul's Church Yard, London.

Mespoulet, M., 2002. From Typical Areas to Random Sampling: Sampling Methods in Russia From 1875 To 1930. Science in Context, 15(3), 411-425.

Middle East Technical University, 2011,
<http://www.stat.metu.edu.tr/people/faculty/oztas>.

Neyman, J., 1934. On the Two Different Aspects of the Representative Method: The Method of Stratified Sampling and the Method of Purposive Selection. *Journal of the Royal Statistical Society*, 97(4), 558–625.

Neyman, J., 1938. Contribution to the Theory of Sampling Human Populations. *American Statistical Association*, 33(201), 101–116.

Pfeffermann D., Rao. C. R., 2009. *Handbook of Statistics 29*, Elsevier, Hungary.

Seneta, E., 1985. A Sketch of the History of Survey Sampling in Russia. *Journal of the Royal Statistical Society*, 148(2), 118–125.

Seng, Y. P., 1951. Historical Survey of the Development of Sampling Theories and Practice. *Journal of the Royal Statistical Society, Series A*, 114(2), 214–231.

Som, R. K., 1998. P. V. Sukhatme’s Contributions to Survey Sampling. *Journal of the Indian Society of Agricultural Statistics*, 51(2&3), 257–264.

The Encyclopedia Sponsored by Statistics and Probability Societies, 2011, Statprob.
<http://statprob.com/encyclopedia/AndersNicolaiKIAER.html>.

Stephan, F. F., 1948. History of the Uses of Modern Sampling Procedures. *Journal of the American Statistical Association*, 43(241), 12–39.

Thomsen, I., Zhang, L. C., 2012. A Prediction Approach to Representative Sampling,
<http://www.stats.gov.cn>.

Zarkovich, S. S., 1962. A Supplement to “Note on the History of Sampling Methods in Russia”, *Journal of the Royal Statistical Society, Series A.*, 125(4), 580-582.

Zarkovich, S.S., 1956. Note on the History of Sampling Methods in Russia. *Journal of the Royal Statistical Society, Series A.*, 119(3), 336-338.

THE HISTORY OF SURVEY SAMPLING

ABSTRACT

Survey sampling might have different roots, which include various activities in fields such as agriculture, public administration and social studies. Collection of large data sets for public administration and social studies provided great opportunities for the improvement of sampling methods. In this study, the historical progress of survey sampling that leads us to random sampling methods are summarized. Although sampling theory and applications constitute a separate and important branch in statistical sciences, it was definitely not the case in the beginning. The roots of survey sampling lie in official statistics and social statistics, rather than probability theory and experimental design. Especially political arithmetic is the primary activity that leads sampling to modern sampling theory. However, sampling is counted as an important branch of statistical science after probability theory has become an important element of sampling theory. In this study, the period between social statistics and modern sampling theory, especially 1895-1934 era and Kiaer, Bowley, and Neyman milestones, are examined in detail.

Keywords: Purposively sampling, Survey sampling, Random sampling, Historical progress, Representative method.