BERNOULLİ TECRÜBESİ ÜZERİNE BİR TÜKETİCİ VE REKLÂM ARAŞTIRMASI

Doç. Dr. M. Kemal YOĞURTÇUGİL

I.

İlmin gayesi benzer problemlerin çözümlenmesinde kullanabilecek genel metodlar tesis etmek olduğuna göre, konusu «çok sayıda tekrarlanma özelliğine sahip olaylar alanında bu tür genel metodlar bulma» şeklinde tanımlıyabileceğimiz ilme, ihtimaller teorisi denir. İhtimaller teorisi günümüz çalışmalarında önemli bir yer işgal eder. Son yüzyıl içerisinde ortaya atılan Einstein'in İzafiyet Teorisi ile Planck'm Quantum Teorisi sadece fizik alanında etki yapmakla kalmamış, yerleşmiş bazı büyük problemleri beklenmedik şekil altında ileri sürerek yeni bir mantığın «İhtimaller Mantığı» ortaya çıkmasına sebep olmuştur. Bu ihtimaller mantığında, doğru - yanlış dediğimiz bir çift süreksiz değer yerine bir ile sıfır arasında değişen ve ençok muhtemelden enaz muhtemele doğru sıralanan bir ihtimaller derecesi kabul edilmektedir.

Bir (A) olayının meydana gelebildiği veya gelemediği bir tecrübe yapılıyor ve bu olayın vuk'u bulma ihtimali *(p)* ye eşit kabul ediliyorsa, deneyin bir çok defalar tekrarlanması ve bu esnada her deneyin sonucunun kendinden sonrakilerin sonucuna tesir etmemesi şartı ile, büyük bir deney sayısı için (A) olayının frekansı, pratik kesinlikle aşağı yukarı, bu olayın *(p)* ihtimaline eşittir. Bu netice ihtimaller hesabının temel kanunlarından biri olarak bilinen Büyük Sayılar Kanunu'nun en basit şeklini ifade eden Bernoulli teoremini teşkil eder.

Böylece Büyük Sayılar Kanunu, çok sayıda yapılan seri deneylerde bir (A) olayının ihtimali bu (A) olayının vuk'u bulma frekansına yaklaşır, ifadesinden hareketle bu ihtimali tecrübî olarak belirtmeyi mümkün kılar. Bir çok hallerde bu tecrübî yol, ihtimalin hesabı için mümkün olan tek çaredir ve ayrıca ihtimal ile frekans arasındaki bu bağlantının bilinmesi, ihtimaller hesabının birçok ilimlerdeki tatbikatının temelini teşkil eder.

Bir fabrikada verilen ve değişmeyen şartlar altında imal edilen mamûllerin % 1.6 sının istenilen vasıfta olmadığını kabul edelim. Bu 1000 mamûlde 16 sının iskartaya çıkarılması demektir. Tabiatıyla bu miktar her 1000 parçalık mamûl yığınları için aynı sayıda olmayacak, fena parçaların sayısı bazan biraz daha fazla veya biraz daha az bulunacaktır; fakat değişmeyen imal şartları altında yapılan çok sayıda imalâtta bu yüzde oranı ortalama olarak aynı kalacaktır. Bu tip ortalama, yani göz önüne alman olayın ihtimali, sadece olmuş hadiseler hakkında bir fikir vermekle kalmaz aynı zamanda olacak hadiselerin neticelerini de tahmin etmeye hizmet ederler.

Olayların iki mümkün durumdan biri veya diğeri şeklinde ortaya çıkmasının zorunlu olduğu hallerde, durumlardan birini uygun olan veya başarı, diğerini ise uygun olmayan veya başarısızlık şeklinde ele alabiliriz. Bir tavla zarımn tek atışta 6 gelmesinin başarı kabul edildiği bir tecrübede başarı ihtimali (1/6); başarısızlık ihtimali ise geride beş yüzün kalması dolayısıyla (5/6) dır. Benzer şekilde hilesiz madeni bir paranın atılışında yazı gelmesi isteniliyorsa başarı ihtimali (1/2); başarısızlı kihtimali de (1/2) dir. Görülüyor ki burada başarı (p), istenilen bir sonucun ortaya çıkması yolları sayısının ortaya çıkması mümkün olayların toplam sayısına bölünmesi suretiyle elde edilmekte ve başarı ile başarısızlık ihtimallerinin toplamı $(p + q = 1)$ olmaktadır.

Biri başarı diğeri başarısızlık olmak üzere sadece iki sonuç veren bir hadise, aynı şartlar altında ve her seferinde elde edilen neticeler yekdiğerinden bağımsız olacak şekilde (n) defa tekrarlansa «Bernoulli Tecrübesi», (p) bir atıştaki başarı ihtimalini temsil ediyorsa ilk (x) tecrübenin başarı olması ihtimali $(p \,.\, p \,.\, p \,\ldots\, p) = p^x$ ile belirir. Geri kalan $(n\text{-}x)$ hadise için başarısızlık aynı şekilde, (q) tek bir artıştaki başarısızlık ihtimalini vermek şartıyla, q^{n-x} dir. İki sonucun, başarı veya başarısızlığın, birlikte vuk'u bulması ihtimali bu iki ihtimalin çarpımına eşit olduğundan $(p^x \,.\, q^{n-x})$ yazılır. Ancak bu ihtimal tek bir sıralanışı vereceğinden, (n) tecrübedeki (x) başarı sayısı ile $(n\text{-}x)$ sayıdaki başarısızlık sayısını içinde bulunduran çeşitli mümkün tertiplerin miktarı $\frac{n!}{x!(n-x)!}$ den hesaplanacak ve neticede iki muhtemel sonuçlu bir hadisenin (n) defa tekerrüründe (x) kere başarı vermesi ihtimali $b(x) = \frac{n!}{x!(n-x)!} \cdot p^x q^{n-x}$ ile belirecektir. Bu ihtimal kanununa, çeşitli değerler için bu kanunun ifadesi Nevton binomunun birbirini takip eden terimlerine karşılık geldiği için, Binom Kanunu adı verilmiştir. Bu şekilde $x=n$ den $x=0$ a kadar değerler için hesaplanacak $b(x)$'leri Y eksenine ve (x)'leri X eksenine yerleştirmek suretiyle elde edilecek grafik (n) tecrübe için muhtelif başarı ihtimallerinin bölünmesini verecektir.

II.

A. Denemenin amacı :

Bu deneme halen iç piyasaya sofra margarini, yemeklik margarin, banyo sabunu, temizleme tozu .... gibi çeşitli mamûlleri arz eden bir büyük şirketin pazarlama servisinde hazırlanmış ve değerlendirilmiştir.

Bu çalışmanın amacı olarak şirket mamûlü (A) marka banyo sabunu için yapılan pazarlama faaliyetlerinin yeterli olup olmadığının tahkiki ve denenecek hipotezlere göre varılacak hükümlerden hareketle talep araştırması, mamûl plânlaması gibi problemlerin çözümlenmesi düşünülmüştür.

Bilindiği gibi istatistik metodlarınm gayesi üzerinde çalışılan ana kütle hakkında bir fikir edinmek ve onun özelliklerini tanımaktır. Ancak bu topluluk içindeki bütün birimleri ele geçirip incelemek ekseriya imkânsız ve çoğu kere lüzumsuz olduğundan ona ait değerler bu yığından alınan bir örnekten hareketle tahmin edilirler. Tecrübenin şekli kullanılacak metodları tayin eder ve istatistikçi elde edilen verileri örnekleme plânına uygun metodlarla işler, neticeleri araştırma sahasına ait bilgilerine dayanarak izah ve tefsir eder. Bu bakımdan aşağıdaki çalışmamızda biri evet diğeri hayır olmak üzere sadece iki muhtemel sonuç veren, aynı şartlar altında ve her seferinde elde edilen neticeler yekdiğerinden bağımsız olacak şekilde (n) defa tekrarlanan, bir hadiseye ait verilerin yorumlanmasında kullanılan örnekleme prensiplerinden istifade edilecektir.

B. *Denemenin kapsamı ve uygulanışı :*

Bu araştırmada çözüm iki safhalıdır :

a. Toplulukta herhangi bir karakteri taşıyanların sayısını veya oranını bulabilmek için bu yığından çekilen örneklerin müşahede edilmesi ve kütleye ait parametrelerin yerini tutacak örnek değerlerinin hesaplanması.

b. Bu değerler esas alınarak yığma ait gerçek değerin tahmini ve bu tahmindeki geçerlilik sınırlarının diğer bir ifade ile genelleme ile yapılabilecek hataların derecesinin belirtilmesi.

Örnekleme tekniğinin en ilgi çekici yönü elâstiki olmasıdır. Örnekleme sayesinde belirli bir örnek üzerinde yürütülen istatistik araştırma, çoğu zaman tam sayım usulüyle yapılan araştırmalarla elde edilmesi maddeten imkânsız olan bilgilerin sağlanmasına imkân vermekte; sonuçlara varan zamanın kısalması, gerçekleştirme maliyetinin düşmesi gibi durumlara yol açmaktadır. Bununla beraber bu metodun sınırlarını ve zorunluluklarını da gözden uzak tutmamak gerekir. Örnekleme yoluyla elde edilen sonuçların kesin olup olmaması, ele alınan örneğin tesadüfi olarak seçilip seçilmediğine ve büyüklüğüne bağlıdır.

Örnekleme ile yapılacak araştırmalarda gerek örnek seçme işlemi gerek soru cedvellerinin hazırlanması ve gerekse anket personelinin yetiştirilmesi ve değerlendirme metodlarının tayini göz önünde tutulması gereken başlıca problemlerdendir. Genel olarak örneğin büyüklüğü incelenen konuya bağlıdır. Müşahede güçlükleri arttıkça örnekleme hatalarından daha sakıncalı olan gözlem hatalarını azaltmak için örnek daha derinlemesine incelenmeli ve daha dar bir çerçevede tutulmalıdır. Diğer taraftan örneğin çekim şekli de ele alınacak büyüklüğe etki eder.

Tabakalaştırma, örnekleme yolu ile yapılan araştırmalara büyük ölçüde kesinlik katmaktadır. Tabakalaştırmanın faydası araştırmada ölçülecek büyüklüklerin tabakalaştırma kıstasları ile sıkı bir bağlantısı olduğu ölçüde artar[1]. Bununla beraber sınırlı bir coğrafi bölgenin,

---

1) G. Calot (Terc. E. Güntekin) «Örnekleme Metodlarının Meskenle İlgili Problemlere Uygulanması» Ankara 1969 sf. 53.

özellikle mahalli problemlerin, analizi çoğu zaman kolay olmamakta, tabakalaştırma usulleri ise şüphesiz bazı sakıncaları ortadan kaldırmakla beraber herzaman uygulanamamaktadır. Diğer bir ifade ile, örnekleme sırasında hane halkı, mesken birimi ve fert sayısı araştırılıyorsa meskenler üzerinde tabakalaştırılmamış sistematik bir çekime baş vurmak yararlı olur. Bu durumda örnek olarak çekilen binalar, bu binalarda bulunan mesken ve içinde oturanlar örnekleme malzemesini meydana getirirler. Örnekleme yoluyla yapılan araştırmaların etkilerini arttırmak için kullanılan tabakalaştırma ve devamlı örneklerden yararlanma teknikleri tam bir sayım yapılamadığı hallerde bazı güçlükler yaratırlar. Diğer taraftan her örnekleme önceden tam sayım ile elde edilmiş bir örnekleme temeline «çerçeve» ihtiyaç gösterir. Bir ülkenin idari düzenine, bazı kayıtların bulunup bulunmamasına ve sayım neticelerinin döküm şekline göre çeşitli çerçeveler kullanılabilir.

Denemenin kapsamı İstanbul Belediyesi hudutları içindeki bütün sosyo-ekonomik guruplardır. Örnek birimi «hane halkı»[2] dır. Örneğin temsili olabilmesi için örnekleme plânı hazırlanırken muhtelif hipotezlerin deneneceği örneklerin mümkün olduğu kadar tesadüfi sondajın[3] şartlarının sağlanması suretiyle çekilmesine çalışılmıştır.

Böylece araştırmamızda «ana kütlenin çok büyük olduğu hallerde eğer örnek Basit Tesadüfi Sondaj'ın şartları sağlanmak suretiyle çekilemiyorsa numuneyi tesadüfi seçimin şartlarını yerine getirmek suretiyle alınız[4] ve örnekleme hatasını - tesadüfi bir numune için örnek-

---

2) Hane halkı, aralarında aile bağı bulunsun veya bulunmasın aynı evde veya evin bir bölümünde yaşayan, aynı kazandan yiyen, gelir ve giderlerini ayırmayan ve hane halkı hizmet ve yönetimine katılan bir veya birkaç kişinin oluşturduğu topluluktur.

3) İstatistik tekniğinde yığını temsil edecek örneğin seçimi ilk ve önemli kısmı teşkil eder; çünkü bu örnekten elde edilecek sonuçların geçerliliği :

a. Örneğin bir ön yargıya yer vermeyecek ve sistematik bir farklılık yaratmıyacak şekilde seçilmesine,

b. Örneğe dahil birimlerden herbirinin yekdiğerinden bağımsız bulunmasına,

c. Verilerin seçildiği alanlarla diğer alanlar arasında temel farklılıkların bulunmasına ve

d. Örneğe alınacak verilerin hepsine istisnasız aynı şartların uygulanmasına bağlıdır. Bkz. H. Arkın - R.R. Colton (Terc. S. Kendir) «İstatistik Metodlar» Ankara 1968 sf. 128.

4) Bilindiği gibi B.T.S. sadece Tesadüfi Sondaj olarak tanımlanan metodun özel bir şeklidir. Diğer bir ifade ile her B.T.S. Tesadüfi Sondaj olduğu halde her

lemeden doğacak hatanın en fazla bu numune ile aynı büyüklükte çekilebilecek bir Basit Tesadüfi Sondaj örneğinin hatası kadar bulunabileceği - prensibinden hareketle hesaplayınız»[5] görüşünden istifade imkânı sağlanmış olacaktır.

Araştırmamızda örnek hane halkının seçimi iki safhalı olmuştur. Birinci safhada örnek cadde ve sokaklar, ikinci safhada örnek mesken birimleri seçilmiştir. Örnek cadde ve sokakların seçiminde kütlenin tamamını kavrayacak ve yığın teşkil eden birimlerin hüviyetini ayırd etmeyi sağlayacak bir vasıta «çerçeve» olarak İstanbul Belediyesi içindeki cadde, sokak ve meydanların isimleri ile haritalarını ihtiva eden ve kazalar esasına göre tertip edilmiş bulunan alfabetik liste[6] kullanılmıştır.

Bu listede mevcut 12 kazaya ait 6133 sokaktan sistematik olarak, denenecek hipotezlerin mahiyetine göre farklı miktarda, tesbit edilen örnek hacminin beşte biri sayıda, sokak adı çekilmiştir. Bu sokaklar o bölgeye ait paftada bulunduktan sonra sıra örnek mesken birimlerinin çekimine gelmiştir. Bu seçimde öncelikle yazı ve tura ile sokağın sağ veya sol cephesi tayin edilmiş; bunu takiben o sokakta görevlendirilmiş müşahide ilk müracaat edeceği birimin sokağın sağ veya solundan kaçıncı binada olacağı yazılı olarak bildirilmiştir. Aynı binada birden fazla yerleşme ünitesi varsa bunlardan birinin tercihi; boş, ticari veya mesleki bir amaçla kullanıldığı tesbit edilen meskenlerin sayım dışı bırakılarak müteakip üniteye gidilmesi; soruya cevap alınamaması halinde bir sonrakine başvurulması; ilk müşahedenin tamamlanması halinde o sokakta beş bina biriminin atlanarak beş üniteye uğranması talimatı dışında, malûmatın tesbit ve kaydedilmesi gibi diğer hususlar anketörlerin tecrübeli ve denenmiş birer eleman olmaları nedeniyle şahsî inisiyatiflerine bırakılmıştır.

Çalışma pratik etkililik sebebiyle basit, tesbit edilmesi kolay olan

---

Tesadüfi Sondaj B.T.S. değildir. B.T.S. için gerekli şartlar şöyle özetlenebilir. Ezcümle bir olayın tahakkuk etme ihtimali *(p)*, aną kütleden çekilebilecek *(n)* birimli muhtelif numuneler için aynı olacak ve hatta münferit bir örnek içinde bile bütün tecrübe boyunca sabit kalacaktır. Ayrıca her tecrübe kendinden evvel ve sonrakilerden bağımsız olmalıdır. Bkz. C.E. Weatherburn «Mathematical Statistics» Cambridge University Press 1957 sf. 116.

5) Bu görüş hakkında daha fazla bilgi için Bkz. F. Conway «Sampling - An Introduction for Social Scientist» London 1967 sf. 108.

6) H. Lokmanoğlu «Haritalı Şehir Rehberi - İstanbul» İstanbul 1955.

sorularla sınırlı olmak zorundaydı ve evet veya hayır şeklinde istenen cevapların sadeliği, verilerin tümü yerine tablo şekline sokulan tasnif sonuçlarının gösterilmesini ve bazı sebeplerle verilerin saklı tutulmasını gerektirmiştir.

III.

Firma sevk ve idaresi ülke çapında ve bakkallar nezdinde yaptığı bir anket neticesinde, (A) marka banyo sabununun ev kadınlarının şahsında işgal ettiği mevkiin nisbi ehemmiyetinin 0.65 olduğunu yıllık faaliyet raporunda ileri sürmüştür. Bu iddianın sadece İstanbul Belediyesi hudutları dahilindeki hane halkı ünitelerine göre geçerli olup olmadığı, bunu takiben yapılacak diğer araştırmalara da bir ön adım teşkil etmesi amacıyla, tarafımızdan tahkik edilmek istenmiştir. Bu konuda yığından sistematik olarak 80 sokak seçilmiş ve her sokakta beş hane halkı birimi müşahede edilerek 400 birimli bir örnek üzerinde çalışılmıştır. Neticeler masa başına intikal edince 400 ünitenin 236'smdan «(A) marka banyo sabunu kullanırmısınız?» sualine evet cevabı alındığı görülmüştür. Ana kütle değeri (P)'nin muayyen bir değer olduğu hipotezinin tahkiki için:

$$\nu^2 = \frac{(a - rb)^2}{r(a+b)}$$

formülü kullanılmıştır[7]. Burada *(a)* örnekteki başarı sayısını, *(b)* başarısızlık sayısını ve *(r)*, *(a/b)* nisbetinin matematik ümidini temsil ettiğine göre

$$\nu^2 = \frac{\left(236 - \frac{0.65}{0.35} . 164\right)^2}{\frac{0.65}{0.35} \; 400}$$

7. G.W. Snedecor «Statistical Methods» The Iowa State College Press 1950 sf. 26.

$$\chi^2 = \frac{(236-(1.8571)(164))^2}{(1.8571)(400)} = 6.22$$

değeri $\chi^2_{0.05}$ değeri ile mukayese edilmiş ve $6.22 > 3.841$ çıktığından ana kütle bölünüşünün 0.65/0.35 veya ana kütle nisbetinin 0.65 olduğu hipotezi kabul edilememiştir.

b. Örnek değeri ($p=236/400$)'nin tesadüfen meydana gelip gelmediği veya $p-P=0.06$'nın anlamlı olup olmadığı hipotezinin testi $\alpha=0.05$ ihtimal eşiği içinde yapılmıştır.

$$H_0 : P = 0.65$$

$$H_1 : P \neq 0.65$$

$$z = \frac{p-P}{\sqrt{\dfrac{P(1-P)}{n}}} = \frac{0.59-0.65}{\sqrt{\dfrac{0.65\times 0.35}{400}}} = -2.52$$

Hesabi z değeri tablo z değeri 1.96 dan mutlak değerce büyük çıktığı için $P=0.65$ hipotezi red edilerek farkın anlamlı olduğu kabul edilmiştir. Buna göre İstanbul şehir nüfusu içinde (A) marka sabunu kullanan ev kadınlarının nisbeti 0.65'den farklıdır.

1. $P=0.50$ ise bu taktirde küçük numuneler için bile örnek değerleri bölünmesi normal bölünmeye yaklaşır. Bu durumda binom bölünme yerine (n)'in büyüyen değerleri için yaklaşımı kabul edilen normal bölünmeyi uygulayabiliriz. Yaklaşım metodu diyebileceğimiz bu usul neticelerin süratle elde edilmesi bakımından çoğu kere esas metoda tercih edilmekte, ancak bu kullanım örnek büyüklüğü ve ihtimal eşiğine göre şartlandırılmaktadır.

2. $P\neq 0.50$ haller için binom bölünme yerine yaklaşımı normal bölünmeyi kullanmak ancak (P); (n) ve (α) verilerine bağlıdır. Gerçekten (n)'in belli bir değeri için ana kütle nisbeti (P); 0.50'ye yaklaştıkça örnek nisbetleri bölünmesi normale yaklaşmakta, aksi halde normal eğri cetvelinden istifade edebilmek için numune mevcudunun büyük tutulması gerekmektedir. Meselâ; araştırmamızda beliren örnek

değeri $p=0.59$; $n=400$ yerine $n=1000$ birimli bir numuneden elde edilmiş olsaydı $P=0.65$ hipotezini red edemiyecektik. Aynı şekilde $n=400$ için $\alpha=0.01$ olsaydı tablo değeri $z=2.58$ olduğundan hipotezin reddi veya kabulü hususunda ihtiyatlı davranmamız gerekecekti.

Tatbikatta, genellikle, normal ihtimal bölünmesinin binom bölünme yerine kullanılabilmesi, (nP) ve $n(1-P)$ nin 5 den «bazı müelliflerce 10»[8] büyük olması şartına bağlanmakta ve ayrıca $P\neq0.50$ için Yates tashihi de kullanılmamaktadır[9].

c. Yukarıda yapılan ve ana kütle oranının 0.65 olduğu hipotezinin reddi ile sonuçlanan iki deneme bu kez başka bir metodla kontrol edilmek istenmiştir. Bu usul ana kütleden

$$n = \frac{r-\chi^2}{[(1+r)k]^2} \qquad ^{10}$$

sayıda çekilen örnekte evet diyenlerin miktarının hayır diyenlerin miktarını numune mevcudunun tesbit edilen bir (k) yüzdesi kadar veya daha fazla aşıp aşmadığının testine dayanmaktadır. Formülde (r) test edilen ana kütle bölünüşünü, (χ^2) bir serbestlik derecesinde istenilen ihtimale göre tabloda bulunan değeri göstermekte; evetlerin sayısı hayırların sayısını örnek mevcudunun tesbit edilen bir (k) yüzdesi kadar veya daha fazla aşması halinde hipotez red edilmektedir. Bu sebepten 400 birimlik örneğimizin içinden çekildiği ana kütlenin evet/hayır şeklindeki bölünüşünün 0.65/0.35 olduğu iddiasının $\alpha=0.05$ seviyesinde yeniden testine girişilmiş, hipotezin evetlerin sayısı hayırların sayısından örnek mevcudunun 0.05 i kadar veya daha fazla çıkması halinde red edilmesi kabul edilmiştir.

Diğer bir ifade ile ana kütleden

$$n = \frac{\left(\frac{0.65}{0.35}\right)(3.841)}{\left[\left(1+\frac{0.65}{0.35}\right)(0.05)\right]^2}$$

$$n = 359$$

8) J.G. Peatman «Introduction to Applied Statistics» London 1963 sf. 263.
9) F.E. Croxton - D.J. Cowden «Applied General Statistics» London 1955 sf. 671.
10) G.W. Snedecor a.g.e. sf. 25.

birim ihtiva eden yeni bir örnek çekilecek ve tecrübe sonucu evetlerin sayısı hayırların sayısını $359 \times 0.05 = 18$ ve daha fazla aşması halinde hipotez red edilecektir.

Gerçekten yeniden hazırlanan bir çekim şeması içinde sistematik olarak 72 sokak seçilmiş ve her sokakta beş mesken birimi müşahede edilerek evetlerin sayısının 227 olduğu görülmüştür. Bu durumda $227 - 133 = 94$ farkı 18 den büyük çıktığı için ana kütle bölünüşünün 0.65/0.35 veya ana kütle nisbetinin 0.65 olduğu görüşü kabul edilememiştir.

Her üç sonuçta bizi aynı hükme götürünce neticede bu örneklerin içinden çekildiği ana kütle nisbeti (P) için güven sınırları tesbitine çalışılmış, evvelâ kaba bir usul olan ve (P)'nin meçhul bir değer olması halinde onun tahmini olarak (p)'nin alınması faraziyesine dayanan metod kullanılmıştır.

$$P_{1,2} = \frac{p + \dfrac{z^2}{2n} \mp z\sqrt{\dfrac{p(1-p)}{n} + \dfrac{z^2}{4n^2}}}{1 + \dfrac{z^2}{n}} \qquad 11$$

formulünden hareket ve (n)'in çok büyük olması halinde ($z^2/2n$); ($z^2/4n^2$) ve (z^2/n) çok küçük olacaklarından ihmal edilerek

a.

$$P_{1,2} = p \mp z \sqrt{\frac{p(1-p)}{n}}$$

$$P_{1,2} = 0.59 \mp 1.96 \sqrt{\frac{0.59 \times 0.41}{400}}$$

$$P_1 = 0.542$$

$$P_2 = 0.638 \quad \text{dir.}$$

11) M.R. Spiegel «Theory and Problems of Statistics» Schaum 1961 sf. 163.

Ancak bulunan bu sınırların ana kütle nisbeti (P)'nin 0.59 ve dolayısıyla standart hatanın 0.0245 olduğu görüşüne göre hesaplandığı göz önünde tutulursa gerçek değerleri aksettirmediği ileri sürülebilir.

Kullanılacak standart hatanın değeri test edilecek hipoteze bağlı olduğundan, bu iddiayı kabul etmek ve gerçek oranın daha güvenilir hudut değerini bulmak mecburiyetindeyiz. Biliyoruz ki, bir güven aralığı bulmak demek istenen güvenlik seviyesinde kabul edilebilir hipotezleri sınırlandırmak demektir. Diğer bir ifade ile, gerçek oranın ele alınan ihtimal seviyesinde kabul edilebilir en küçük ve en büyük hipotetik değerleri hesaplanmalıdır. Bu değerler

$$\frac{p-P}{\sqrt{\dfrac{P(1-P)}{n}}} = z$$

eşitliğinde (z); (p) ve (n) yerine değerlerini koymak suretiyle (P)'ye göre tanzim edilecek bir ikinci derece denkleminin kökleri olacaktır. Bu kıymetler daha süratle yukardaki denklemin köklerini veren

$$P_{1,2} = \frac{(2p + \dfrac{z^2}{n}) \mp \sqrt{(2p + \dfrac{z^2}{n})^2 - 4p^2 (1 + \dfrac{z^2}{n})}}{2(1 + \dfrac{z^2}{n})}$$

eşitliğinden elde edilebilir. Bu şekilde uygulama binom bölünme için geçerli olup eğer ana kütle hipergeometrik ise, eşitlikteki her (z^2/n) yerine $(z^2/n)(N-n/N-1)$ konulmalıdır. Ancak bu metod büyük örneklerde kullanılmakta (p)'nin 0.50 ye yakın değerleri için $(0.7 > p > 0.3)$ daha ziyade numune değerinden hareketle hesaplanan standart hatadan istifade eden kaba yol işlem kolaylığı bakımından tercih edilmektedir.

Gerçek oranın 0.05 seviyesinde kabul edilebilir en büyük ve en küçük hipotetik değerleri b.

$$\frac{P-0.59}{\sqrt{\frac{P(1-P)}{400}}} = 1.96 \quad \text{veya}$$

$$403.84\ P^2-475.84\ P+139.24=0$$

denkleminin kökleri olan

$$P_{1,2} = \frac{\frac{475.84}{400} \mp \sqrt{(\frac{475.84}{400})^2 - 1.3924\ (\frac{403.84}{400})}}{2(\frac{403.84}{400})}$$

$P_1=5.40$ ce $P_2=0.638$ kıymetleridir.

Ana kütle nisbetleri için iki farklı şekilde hesapladığımız güven sınırları ayrıca üçüncü bir yoldan da tahkik edilmiştir. Bu defa

c.

$$\nu^2 = \frac{[x - \frac{P}{1-P}(n-x)]^2}{\frac{P}{1-P}.n}$$

formülü kullanılmış ve

$$3.841 = \frac{[236 - \frac{P}{1-P}(400-236)]^2}{\frac{P}{1-P}.400}$$

$$161.536.4\ P^2-190.336,4\ P+55.696=0$$

denkleminden hareketle

$$P_2=0.637 \text{ ve } P_1=0.541$$

bulunmuştur[12].

Çalışmamızın buraya kadar olan kısmında biri 400 diğeri 360 olmak üzere birbirinden bağımsız iki örnek müşahede etmiş ve numune değerinden hareketle ana kütle nisbetinin 0.64 ile 0.54 değerleri arasında olduğunu hesaplamış bulunuyoruz. Bu sonuca göre (A) marka sabunun ev kadınları nezdinde işgal ettiği mevkiin nisbi ehemmiyetinin firmanın iddia ettiği gibi 0.65 olmayıp daha az olduğu söylenmiştir. Buna karşılık sevkî idare iddiasında israr etmiş, beliren bu anlaşmazlığın giderilmesi için üçüncü bir örneğin çekiminde karar kılınmıştır. Bu kez örnek zümrelere göre sondaj usulüne göre çekilecek ve ayrıca mevcudu da arttırılacaktır.

Yapılan yeni bir örnekleme plânına göre İstanbul Belediyesi hudutları içinde yaşayan[13] 1.570.505 (1965) kişi tartıları sırasıyle $W_A=0.41$; $W_B=0.37$ ve $W_C=0.22$ olan ve A,B,C diye adlandırılan üç alt bölgede toplanmıştır.

A	B	C
Fatih	Beyoğlu	Kadıköy
Eminönü	Şişli	Üsküdar
Eyüp	Beşiktaş	Beykoz
Bakırköy	Sarıyer	Adalar

Örnek mevcudu dört misli büyültülerek 1600 birime çıkarılmış ve bölgeler itibariyle dağılım ise tartılara dayanılarak $n_A=656$; $n_B=592$ ve $n_C=352$ sayılarında tesbit edilmiştir. Her bir alt bölgede örneğe girecek birimlerin yeri yine sistematik olarak çekilen cadde ve sokaklara, daha önce de izah edildiği şekilde, dağıtılmış ve elde edilen veriler tasnif sonucu aşağıdaki şekilde özetlenmiştir.

---

12) P_1 ve P_2 kıymetlerinde bir evvelki metoda göre beliren küçük farklar yuvarlaklaştırma hatalarıdır.

13) 1965 yılında İstanbul Belediyesi hudutları içinde 14 kazada toplam 1.742.978 kişi bulunmaktaydı. Ancak çerçeve olarak kullandığımız listenin tanzinii sırasında Gaziosmanpaşa ve Zeytinburnu kazaları olmadığından bu iki kaza araştırma dışı bırakılmıştır.

TABLO : 1

Bölgeler	*Evet*	*Hayır*	*Toplam*
A	437	219	656
B	287	305	592
C	195	157	352
Toplam	919	681	1600

Bu tablodan ve

$$W_A = 0.41 \qquad p_A = 0.67$$

$$W_B = 0.37 \qquad p_B = 0.48$$

$$W_C = 0.22 \qquad p_C = 0.55$$

verilerinden hareketle ana kütle nisbetinin tahmini olarak

$$\begin{aligned} P = \sum W_h \,.\, p_h &= (0.41)\,(0.67) + (0.37)\,(0.48) + (0.22)\,(0.55) \\ &= 0.2747 + 0.1776 + 0.1210 \\ &= 0.5733 \end{aligned}$$

bulunmuştur. Bu tahminin variyansı

$$\mathrm{Var}(p_{nis}) = \sum W_h^2 \, \frac{p_h(1-p_h)}{n_h}$$

formülünü kullanarak

$$\mathrm{Var}(p_{nis}) = \left[\, (0.41)^2 \left(\frac{67 \times 33}{656}\right) + (0.37)^2 \left(\frac{48 \times 52}{592}\right) + (0.22)^2 \left(\frac{55 \times 45}{352}\right) \right] 10^{-4}$$

$$= 1.48 \times 10^{-4} \quad \text{dür.}$$

Şimdi elimizde biri 400 diğeri 1600 birim ihtiva eden iki örnek ile iki nisbet ($p_{400} = 0.59$; $p_{1600} = 0.57$) mevcuttur. Biz derhal bu iki örnek değeri arasındaki farkın anlamlılığı üzerinde durmak istedik. Ancak böyle bir teste girişmeden önce zümrelere göre sondajın bize sağladığı kazancı tesbite çalıştık ve gördük ki bu 1600 birimlik örnek basit tesadüfi sondajla çekilmiş olsaydı.

$$p_0 = \frac{919}{1600} = 0.5744 \quad \text{ve}$$

$$\text{Var}(p_0) = \frac{0.5744 \times 0.4256}{1600} = 1.53 \times 10^{-4}$$

bulunacaktı. Elde edilen bu değerlere ve

$$\frac{\text{Var}(p_{nis})}{\text{Var}(p_0)} = \frac{1.48}{1.53} = 0.987$$

sonucuna göre örneğin zümrelere göre sondaj usulü ile çekilmesinin bize büyük bir kazanç temin etmediğini söyleyebileceğiz. Gerçekten aranan vasfı haiz olanların ana kütle içindeki nisbetini zümrelere göre sondaj usulü ile tahmin etmek istediğimizde tahminin sıhhat derecesini ancak onbinde iki nisbetinde arttırabilmiş olduk. Bunda zümre nisbetlerinin $p_A=0.67$; $p_B=0.48$ ve $p_C=0.55$ gibi birbirine çok yakın çıkmasının büyük ölçüde tesiri olmuştur. Böylece «Zümrelere Göre Sondajda Basit Tesadüfi Sondaja nazaran bir kazanç elde edebilme, zümre nisbetlerinin birbirinden çok farklı çıkacak şekilde homojen zümreler kurabilme imkânına bağlıdır» prensibinin doğruluğu bir defa daha açıkça belirmektedir. O halde zümre nisbetleri tahmin edilmek istenen ortalama nisbet etrafında toplanıyorlarsa, bu takdirde zümrelere gitmeden sadece numune mevcudunda küçük bir artış yapmak suretiyle aynı kazancı elde etmemiz mümkündür. Bizim örneğimiz için bu kazanç

$$\frac{1.53}{1.48} \times 1600 = 1654 \quad \text{ve}$$

$$1654 - 1600 = 54$$

birimdir. Diğer bir ifade ile, 1654 birimlik basit tesadüfi sondaj örneği ile 1600 birimlik zümrelere göre sondaj örneği aynı sıhhattedir.

Belli bir ana kütleden herhangi bir özelliği esas almak suretiyle iki örnek çekildiğinde, bu örneklerde muayyen bir vasfı haiz olanların nisbetleri arasında sadece bu numunelere giren birimlerin tesadüfi değişmelerine bağlı bir farklılık ortaya çıkabilecektir. Eğer ana kütleden buna benzer şekilde çok sayıda örnek çiftleri çekilse ve bu örnek çiftlerinin nisbetleri arasındaki farklar kaydedilerek bir frekans dağılımı teşkil edilse, elde edilenin bir normal dağılım olduğu görülecektir. Bütün bu örnek çiftlerinin aynı topluluktan çekildiklerine ve tesbit edilecek farklar sadece tesadüfi veya ârızi farklar olabileceklerine göre gerçek farkın sıfır olması gerekir. Diğer bir ifade ile, eğer pek çok sayıda örnek çifti teşkil edilseydi, bu farklar ortalaması sıfır olacaktı. Buna göre sonsuz sayıda örnek çifti oranları arasındaki farkların dağılımı normal bir eğri ile ifade edilir. Bu dağılımın ortalaması sıfır olacak, tesadüfe bağlı olarak beliren çeşitli farkların dağılımının standart hatası da

$$\sigma_{p2-p1} = \sqrt{P(1-P)\ (1/n_1+1/n_2)}$$

ile hesaplanacaktır. Bu durumda p_2-p_1 farkı

$$1.96\,.\sqrt{P(1-P)\ (1/n_1+1/n_2)}$$

den büyük çıkarsa $\alpha = 0.05$ için $p_2 > p_1$ diyebiliriz.

Aynı test p_2 ve p_1 değerleri için ana kütle nisbetinin güven sınırlarını ayrı ayrı hesaplamak ve (p_2) ile (p_1) için bulunan bu limitlerin birbiri içine girmesi halinde farkın tesadüfen meydana geldiğini kabul etmek; aksi halde muayyen bir (P) değeri için bu örneklerin içinden çekildiği ana kütlelerin farklı veya farkın anlamlı olduğunu ileri sürmek şeklinde de uygulanabilirdi. Ancak bu tür bir testin manâlı sayılabilecek farkları red ettiğine dayanarak makbul bir test olmadığı söylenebilir. Gerçekten bazı problemler için görülmüştür ki, p_2-p_1 farkı manâlı olduğu halde sınırlar sırf üst üste bindiği için farkın anlamlı olduğu hipotezi red edilmektedir. Bu red ediliş tabiatıyle (p_2) açıklığının alt hududuyla, (p_1) açıklığının üst hududunun kesişmesi şansına bağlanmakta, bu ihtimalin (p_2) ve (p_1)'nin ayrı ayrı bu uçlarda bulunabilme şansından daha küçük olduğu kabul edilmektedir[14].

Sınırların kesişmesi yerine tabiidir ki daha elverişli olanı

$$H_0 : p_2 = p_1 = p$$

$$H_1 : p_2 \neq p_1$$

---

14) R.G. Allen «Statistics for Economists» Hutchinson Ltd. London 1966 sf. 193.

hipotezinin test edilmesidir. Ancak bu test de kullanılan standart hatanın hesabında yer alan (P) meçhul bir değer olduğundan onun yerine tahmini kıymeti

$$\bar{p} = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2} = \frac{x_1 + x_2}{n_1 + n_2}$$

alınmakta, ayrıca iki sondaj nisbeti ortalaması %5 den büyükse tashih faktörü olarak

$$\sqrt{1 - \frac{n_1 + n_2}{N_1 + N_2}}$$

yazılmaktadır. Eğer örnekler için başarı nisbetleri yerine başarı sayıları verilmişse bu takdirde

$$\frac{x}{\sigma} = \frac{p_2 - p_1}{\sigma_{p2-p1}}$$

manidarlık oranı yerine büyük numuneler için

$$\frac{x}{\sigma} = \frac{n_2 x_1 - n_1 x_2}{\sqrt{\dfrac{n_1 n_2 (x_1 + x_2)(n_1 + n_2 - x_1 - x_2)}{n_1 + n_2}}}$$ [15]

formülü kullanılmalıdır. Küçük örnekler için normal yaklaşım yukarıdaki eşitliğin payından $(n_1 + n_2)/2$'nin çıkarılmasıyla elde edilir. $n_1 = n_2 = n$ olması halinde yukarıdaki oran

$$\frac{x}{\sigma} = \frac{x_1 - x_2}{\sqrt{\dfrac{(x_1 + x_2)(2n - x_1 - x_2)}{2n}}}$$

15) A.M. Turtle «Elementary Business and Economic Statistics» MacGraw Hill 1957 London sf. 303.

şeklinde hesaplanmalıdır. $\alpha = 0.05$ için bu değerin de 1.96'dan büyük çıkması halinde $p_2=p_1$ hipotezinin red edileceği tabiidir.

Şimdi

$$H_0 : p_2 = p_1 = p$$
$$H_1 : p_2 \neq p_1$$

hipotezinin testine, diğer bir ifade ile $p_1=236/400$ ve $p_2=919/1600$ nisbetleri arasındaki farkın anlamlı olup olmadığının tahkikine girişeceğiz.

$$P = \bar{p} = \frac{(0.59)(400) + (0.57)(1600)}{400+1600} = \frac{1155}{2000} = 0.5775$$

$$\sigma_{p2-p1} = \sqrt{(0.5775)(0.4225)\left(\frac{1}{400} + \frac{1}{1600}\right)}$$

$$= 0.0275$$

çözümlerine dayanılarak $p_2-p_1=0.02$ farkı 0.73 standart hata olduğu ve 1.96 dan küçük çıktığı için $p_2=p_1$ hipotezi red edilememekte, görülen farkın tesadüfen meydana geldiği kabul ve ana kütle nisbetinin 0.65 den kat'iyetle küçük olduğu ileri sürülebilmektedir.

$$H_0 : p_2 = p_1 = p$$

$$H_1 : p_2 \neq p_1$$

testi bu kez

$$\chi^2 = \sum \frac{(x_i - e_i)^2}{e_i}$$

ile hesaplanan örnek değerleri teorik bölünmesinden hareketle uygulanacaktır. Burada her bir örnek için olması lâzım gelen teorik frekanslar (e_i), örnek mevcutları ile ortalama nisbetin çarpımı ile elde edilecektir.

TABLO : 2

	A	$-A$	*Toplam*
Ornek 1	236	164	400
Örnek 2	919	681	1600
Toplam	1155	845	2000

$$e_{11} = 400 \times 0.5775 = 231$$

$$e_{12} = 400 \times 0.4225 = 169$$

$$e_{21} = 1600 \times 0.5775 = 924$$

$$e_{22} = 1600 \times 0.4225 = 676$$

$$\chi^2 = \frac{(236-231)^2}{231} + \frac{(164-169)^2}{169} + \frac{(919-924)^2}{924} + \frac{(681-676)^2}{676} =$$

$$\chi^2 = 0.1082 + 0.1479 + 0.0270 + 0.0369 = 0.32$$

olduğundan teorik ve fiilî frekanslar arası farkların şansa bağlanacağını söyler yani $H_0 : p_2 = p_1$ hipotezini red edemeyiz.

Ayni ki-kare değeri

$$\chi^2 = \frac{(ad-bc)^2(a+b+c+d)}{(a+b)(c+d)(a+c)(b+d)}$$

$$\chi^2 = \frac{(236\times 681-919\times 164)^2\times 2000}{1155\times 845\times 400\times 1600}$$

$$\chi^2 = 0.32$$

şeklinde süratle hesaplanabilir.

Bu tip misaller için χ^2 yi çabuk hesaplama imkânı veren bir diğer formülde

$$\chi^2 = \frac{(n_1+n_2)\,(n_2x_1-n_1x_2)^2}{n_1n_2(x_1+x_2)\;[(n_1+n_2)-(x_1+x_2)]}$$

dir. Burada (x_1); (n_1) örneğindeki ve (x_2); (n_2) örneğindeki başarı sayılarıdır.

Ki-kare testinde (χ^2) kıymeti bir kare değer olduğu için teorik ve gerçek değerler arasındaki farkın yönünü ortadan kaldırır. Böylece ki-kare testi iki taraflı bir z testine benzer. Bununla beraber genellikle frekanslara ait hipotezlerin manâlandırılmasında kullanılan bu testin yine frekanslar için kullanılan z testinden daha fazla tatbik imkânı olduğu söylenecektir; çünkü:

1. z testinin tatbikatı sadece iki muhtemel sonuçlu bölünmelere aittir ve uygulama sahası bu bölünme ile kısıtlıdır[16]. Halbuki tatbikatta öyle tecrübeler olabilir ki muhtemel sonuçları mutlaka iki şekilde değerlendirmek imkânsızlaşır. Meselâ bir firma imal edip sattığı bir malının piyasada tutulup tutulmadığını tahkik için yaptığı bir araştırmada müşterilerinden iyi, orta ve fena değil gibi cevaplar alabilir veya mevcut iki adaydan birine oy verilecek bir toplulukta kararsız olanları da bulunabilir. Bu ve bunun gibi misalleri çoğaltmak mümkündür. Bu gibi durumlarda (χ^2) testi kullanılmalıdır. Bu test genellikle toplam frekansın şıklar arasında eşit dağıtıldığı faraziyesini tahkik şeklinde uygulanır[17].

2. Bir ana kütleden birbirini takiben bağımsız ve tesadüfi olmak üzere çekilen ikiden fazla örnek için aynı vasfı haiz birimlerin nisbeti hesaplanmışsa, bu nisbetler arasındaki farkın anlamlılığı, diğer bir ifade ile

$$H_0 : p_1 = p_2 = p_3 = \ldots = p_k = p$$

$$H_1 : p_1 \neq p_2 \neq p_3 \neq \ldots \neq p_k$$

hipotezinin tahkiki (χ^2) değerinin hesabına dayanacaktır.

	Örnek 1	Örnek 2	...	Örnek k	Top.
Başarı sayısı	n_{11}	n_{12}	...	n_{1k}	$n_{1\cdot}$
Başarısızlık sayısı	n_{21}	n_{22}	...	n_{2k}	$n_{2\cdot}$
Top.	$n_{\cdot 1}$	$n_{\cdot 2}$	...	$n_{\cdot k}$	n

16) R. Loveday «A Second Course in Statistics» Cambridge University Press 1964 sf. 51.

17) J.G. Peatman a.g.e. sf. 252.

Tablodaki n değerleri altındaki işaretlerden birincisi başarı veya başarısızlık vasfını, ikincisi ise örnek sırasını gösterir. Hakiki P nin tahmini

$$\overline{p} = \frac{n_{11}+n_{12}+n_{13}+ \dots +n_{1k}}{n} = \frac{n_{1.}}{n} \quad \text{dir.}$$

Bu değeri teker teker numune birim sayısı ile çarparak başarı sayısı için teorik frekansları, $(1-\overline{p})$ değerini aynı şekilde numune birim sayısı ile çarparak başarısızlık için teorik frekansları bulabiliriz. Bu şekilde elde edilen teorik değerler (e_{ij}) ise

$$\chi^2 = \sum_{i}^{2} \sum_{j}^{k} \frac{(n_{ij}-e_{ij})^2}{e_{ij}}$$

ifadesi ile elde edilen numune değerleri dağılması teorik ki-kare bölünmesine yaklaşır.

a. Teorik frekanslar çok küçük çıkarsa (χ^2) bölünmesi kullanılmamalıdır. Genellikle bu metod (e_{ij})'lerin 5 den büyük[18] olduğu haller için geçerli kabul edilmekte, aksi halde uygulama birbirini takibeden küçük (e_{ij})'li örneklerin kümülasyonu ile başlamaktadır. Bu halde serbestlik derecesi kümülasyonda kaç kutu birleştirilmişse onun bir eksiği kadar azaltılacaktır.

b.

$$\chi^2 = \sum \frac{(n_{ij}-e_{ij})^2}{e_{ij}}$$

bölünmesi aslında kesikli bölünmeler içindir. Sürekli bölünmelere uygulanırken

$$\chi^2 = \sum \frac{\left[\,|n_{ij}-e_{ij}| - \frac{1}{2}\,\right]^2}{e_{ij}}$$

18) Bu değer Fisher'e göre 5; Aitken'e göre 10 ve Kendall'a göre en az 20 olmalıdır.

şeklinde Yates tashihi yapılmalıdır. Genellikle bu tashih serbestlik derecesi bir veya bire yakın (χ^2) değerleri için geçerlidir. Büyük örnekler için tashihli ve tashihsiz (χ^2) aynı sonuçlara yaklaşır. Eğer numune küçükse ve teorik frekansların herbiri 5 ilâ 10 arasında değişiyorsa tashihli ve tashihsiz (χ^2)'ler mukayese edilmeli, test farklı sonuçlar veriyorsa bu takdirde ya örneği büyültmeli veya mültinomial ihtimal bölünmesi tatbik edilmelidir. Biraz önce verilen basit usuller için bu tashih işlemi[19]

$$\chi^2 = \frac{[\,|a-rb| - (r+1)/2]^2}{r(a+b)}$$

$$\nu^2 = \frac{[(ad-bc) - 1/2(a+b+c+d)]^2(a+b+c+d)}{(a+b)(c+d)(a+c)(b+d)}$$

şeklinde uygulanmalıdır.

c. Teorik frekanslar $n_{.j}(n_{1.}/n)$ ile hesaplanmışsa, teorik toplam başarı frekansı fiili toplama eşit çıkacaktır.

d.

$$\chi^2 = \sum \frac{(n_{ij})^2}{e_{ij}} - n$$

yazılış şekli daha ziyade frekansların tam sayı çıkmaması halinde elverişli bir yoldur.

Yukarıda $2 \times k$ tablosu için yapılan genelleştirme $r \times k$ tablosu içinde geçerlidir. Ancak serbestlik derecesi $(k-1)$ yerine $(r-1)(k-1)$ alınmalıdır.

k aynı ana kütleden çekilen örnekleri

r aynı ana kütle içindeki çeşitli karekterleri göstermek üzere

$$\chi^2 = \sum_{i}^{r}\sum_{j}^{k} \frac{(n_{ij}-e_{ij})^2}{e_{ij}}$$

19) G.W. Snedecor a.g.e. sf. 26 ve 199.

ile beliren değişken $(r-1)(k-1)$ serbestlik dereceli bir ki-kare bölünmesi gösterir ve muhtelif örneklerde aynı karekter için eşit kabul edilen nisbetlerin testinde kullanılır.

Şimdi 1600 birimlik örneğimiz üzerinde yeni bir teste girişeceğiz. Bu tahlil (A) marka banyo sabunu kullanma oranı ile yerleşme yeri arasında bir münasebet olup olmadığının araştırılması düşüncesine dayandırılmıştır. Gerçekten veriler gözden geçirilirse A bölgesinde ev kadınlarının 0.67 si (A) sabununu kullanırken, bu nisbet B bölgesi için 0.48 ve C bölgesi için 0.55 dir. Başlangıç hipotezimiz kullanma ile yerleşme yeri arasında bir münasebet olmadığı, yani sabun kullanan ev kadınlarının tesbit edilen bölgeler itibariyle dağılışlarının tesadüfen meydana geldiği olacaktır.

a. Bu konuda evvelâ kullanma nisbetinin bölgeler itibariyle dağılışının aynı olduğu görüşüne göre teorik frekansları hesapladık. Neticeler toplu halde aşağıdadır.

TABLO : 3

Bölgeler	*Fiili frekans* (x_i)	*Teorik frekans* (e_i)	$x_i - e_i$
A	437	306.3	+130.7
B	387	306.3	19.3
C	195	306.3	111.3
Toplam	919	919.0	0.0

$$\chi^2 = \sum \frac{(x_i - e_i)^2}{e_i} = 1.2 + 55.8 + 40.4 = 97.4$$

2 serbestlik derecesinde $\chi^2_{0.05} = 5.99$ olduğundan başlangıç hipotezi red edilmiş ve sonuç olarak yerleşme ile kullanma arasında bir bağ vardır denmiştir.

b. Bu problem için teorik frekanslar ayrıca bağımsızlık faraziyesine göre de hesaplanmış, bu yolda ana kütle nisbetinin tahmini olarak $p = 919/1600 = 0.5744$ alınmıştır.

Teorik frekanslar

$$e_{11} = 656 \times 0.5744 = 377$$
$$e_{21} = 592 \times 0.5744 = 340$$
$$e_{31} = 352 \times 0.5744 = 202$$
$$e_{12} = 656 \times 0.4256 = 279$$
$$e_{22} = 592 \times 0.4256 = 252$$
$$e_{32} = 352 \times 0.4256 = 150$$

Toplam 1600

şeklinde bulunmuştur.

TABLO : 4

	Evet		*Hayır*		
Bölgeler	*Fiili*	*Teorik*	*Fiili*	*Teorik*	*Toplam*
A	437	377	219	279	656
B	287	340	305	252	592
C	195	202	157	150	352
Toplam	919	919	681	681	1600

$$\chi^2 = \frac{(437-377)^2}{377} + \frac{(287-340)^2}{340} + \frac{(195-202)^2}{202} +$$
$$+ \frac{(219-279)^2}{279} + \frac{(305-252)^2}{252} + \frac{(157-150)^2}{150}$$

$$\chi^2 = 9.549 + 8.261 + 0.242 + 12.903 + 11.146 + 0.326$$
$$= 42.427$$

$\chi^2 > \chi^2_{0.05}$ veya $42.427 > 5.99$ olduğundan yine bağımsızlık hipotezi red edilmiş ve neticede yerleşme yeri ile kullanma oranı arasında bir münasebetin varlığı kabul edilmiştir.

Bu sonuç üzerine firma pazarlama servisi B bölgesindeki nisbetin ($p_B = 0.47$) diğer iki bölge nisbetleri ($p_A = 0.67$; $p_C = 0.55$) den daha düşük olması nedenlerini inceleyen ve gelir bölüşümü, aynı firma mamûlü diğer bir marka sabun tercihleri, başka firmalar mamûlleri banyo sabunları satış tahminleri, toz deterjan kullanımındaki artışlar, dağıtım kanallarındaki aksaklıklar ... gibi çeşitli faktörler üzerinde teferruatıyla duran bir raporu müdürler kuruluna götürmüş ve neticede

a. Ana dağıtım deposunun yeniden tesbit edilecek bir yerde kurulmasına ve dağıtım şemasının re-organizasyonu üzerinde çalışmalar yapılmasına,

b. B bölgesindeki nisbî kullanımı arttırmak amacıyla bu bölgede yoğun bir reklâm kampanyasına girişilmesine
karar verilmiştir.

Reklâm kampanyası ile ilgili araştırmada bu kez tesbit edilecek devamlı bir örnek ile çalışılması düşünülmüştür. Gerçekten gelişmeleri «fiyat seviyesi, ifade edilmiş arzulardaki değişmeler gibi...» ölçmek amacıyla yapılan bazı araştırmalarda her seferinde bir örnek alma cihetine gidilmeyip devamlı veya kısmen yenilenen bir örneğin kullanılması faydalı olabilmektedir.

Bu usul :

a. Örnek birimlerinin gösterdikleri niteliklerin anketten ankete karşılaştırılmasına ve dolayısıyla gelişmeler hakkında daha tam bir değerlendirmenin yapılmasına fırsat verir. Ayrı ayrı örnekler sadece nihai bazı farkları açığa çıkarırken, böyle bir yol değişikliklerin karşılaştırmalı bir tablosunu sağlar.

b. İkinci ankette, ilk anket bültenlerinde bulunan bazı soruların tekrarına mani olur.

c. Örnek birimlerinin tesbiti sırasında sayım yetkililerinin işini kolaylaştırır.

Devamlı bir mesken örneği B bölgesini teşkil eden yığından sistematik olarak çekilen 20 cadde ve sokaktaki 5 er hane halkı biriminden meydana gelmiştir. Böylece tesbit edilen 100 hane halkı birimine firma mamûlü (A) marka banyo sabununu kullanıp kullanmadıkları sorulmuş ve 25 evet cevabı alınmıştır. Bunu takiben bölgede yoğun bir reklâm kampanyasına girişilmiş, o bölge cadde ve sokaklarına duvar afişleri yapıştırılmış, bakkallarla bu konuda işbirliği yapılmış ve nihayet bir ay sonra aynı örnek tekrar müşahede edilmiştir. Neticede ikinci ankette ilk müşahedede evet diyen 25 birimin bu defa sadece 5 i evet derken, ilk seferde hayır diyen 75 kişinin 45 inden evet cevabı alınmıştır. Araştırmamızda ikinci ankette, birinci anket sırasına oturanların taşınmış olmaları halinde halihazır oturanlara ilk anketteki soruların tekrar edilmesiyle anket alanına giren kimselerin hafızasına başvurmak gerekmiştir.

TABLO : 5

		Birinci Müşahede		
		Evet	*Hayır*	*Toplam*
İkinci	Evet	5	45	50
Müşahede	Hayır	20	30	50
	Toplam	25	75	100

Tablo verilerine göre kullanma nisbetinde görülen 0.25 bir artışın reklâmdan doğup doğmadığının tahkiki için ki-kare testi uygulanmış; bu testte başlangıç hipotezi olarak görülen farkın tesadüfen meydana geldiği, reklâmın tesiri olmadığı faraziyesi alınmıştır.

Yukarıdaki verileri $\begin{pmatrix} a & c \\ b & d \end{pmatrix}$ şeklinde 2×2 li bir tablo olarak düşündüğümüzde birinci müşahedede evet diyenlerin nisbeti $p_1 = a+b/100 = 0.25$, ikinci müşahedede evet diyenlerin nisbeti $p_2 = a+c/100 = 0.50$ olarak bulunur.

$$p_2 - p_1 = \frac{a+c}{100} - \frac{a+b}{100} = \frac{c-b}{100}$$

ye göre nisbetler farkı c ve b frekansları arasındaki fark olarak belirdiği ve neticede bu tür problemler için ki-kare

$$\chi^2 = \frac{(c-b)^2}{c+b} \quad [20]$$

20) W.G. Cochran «Comparison of Percentages in Matched Samples» Biometrica 37-1950 sf. 256.

şekliyle hesaplandığı için

$$\chi^2 = \frac{(45-20)^2}{45+20} = \frac{625}{65} = 9.61$$

elde edilir. Bu değer $\chi^2_{0.05}$ den büyük olduğu için müstakillik faraziyesi yani birinci müşahededeki evetler ile ikinci müşahededekiler arasında bağımsızlık vardır, hipotezi red ve dolayısıyla reklâmın tesiri olduğu kabul edilmiştir.

Bir serbestlik derecesinde $z=\sqrt{\chi^2}$ özelliğinden $z=3.1$ dir. $\alpha = 0.05$ için bu değer tablo değerinden büyük çıktığı için aym sonuca varılır ve 0.25 farkın manâlı olduğu kabul edilir.

Araştırma servisi reklâm kampanyası ile ilgili çalışma devam ederken 1600 örnekten elde edilen 919 evet'in ikili bir tasnifine girişmiştir. Bu ayırım aşağıda görüleceği gibi bölgeler ve hane tiplerine göre yapılmıştır. Bölgeler daha önceki testde tariflehmiş zümreleri temsil ederken, hane tiplerinden *(a)*; aile meydana getirmeyen haneleri (tek kişilik haneler dahil); *(b)*, tek aileli haneleri ve *(c)* ise çok aileli haneleri tanımlamaktadır.

Tablo verilerine göre A bölgesindeki 438 evet'in 158 adedi aile meydana getirmeyen hanelerden; 172 si tek aileli hanelerden ve 107 si ise çok aileli hanelerden alınmıştır. Buna karşılık C bölgesinde bu değerler sırasıyla 42; 117 ve 36 dır. Toplam olarak 919 evet'in 297 si *(a)* tipi hanelere; 397 si *(b)* tipi hanelere ve 228 i ise *(c)* tipi hanelere aittir.

TABLO : 6

Bölgeler	*Hane Tipleri* *(a)*	*(b)*	*(c)*	*Toplam*
A	158	172	107	437
B	97	105	85	287
C	42	117	36	195
Toplam	297	394	228	919

İkili kritere göre tanzim edilmiş bir tablo için variyans analizinde kullanılacak formüller aşağıda verilmiştir. Yalnız bu hesapların, bir gurup kıymetin variyansı onun orijininden müstakil olduğu için, 100 orijinine göre yapıldığı ve sıra ve sütunlar arasında karşılıklı bir tesirin olmadığı faraziyesinin kabulü ile geçerli olduğu[21] unutulmamalıdır.

$$T = \sum\sum x_{ij} \qquad T_i = \sum x_{ij}$$

$$\sum\sum (x_{ij} - \bar{x})^2 = \sum\sum x^2_{ij} - N\bar{x}^2 = \sum\sum x^2_{ij} - \frac{T^2}{N}$$

$$\sum\sum (x_{ij} - \bar{x}_i)^2 = \sum\sum x^2_{ij} - \left(\frac{T_i^2}{n_i}\right)$$

$$\sum (\bar{x}_i - \bar{x})^2 . n_i = \sum \left(\frac{T_i^2}{n_i}\right) - \frac{T^2}{N}$$

TABLO : 7

Hane Tipleri

Bölgeler	*(a)*	*(b)*	*(c)*	T_i	T^2_i	x^2_{ij}
A	58	72	7	137	18769	8597
B	—3	5	—15	—13	169	259
C	—58	17	—64	—105	11025	7749
T_j	—3	94	—72	19	29983	16605
T^2_j	9	8836	5184	14029		
x^2_{ij}	6737	5498	4370	16605		

Tablo verilerine göre:

a. Sapmaların kareleri toplamı

---

21) R. Goodman «Statistics» The English Üniversities Press London 1960 sf. 177.

$$\sum\sum x^2_{ij} - \frac{T^2}{N} = 16605 - \frac{19^2}{9} = 16565$$

b. Hane tipleri için kareler toplamı

$$\sum\left(\frac{T_{j.}^2}{n_j}\right) - \frac{T^2}{N} = \frac{14029}{3} - \frac{19^2}{9} = 4636$$

c. Bölgeler için kareler toplamı

$$\sum\left(\frac{T_{i.}^2}{n_i}\right) - \frac{T^2}{N} = \frac{29963}{3} - \frac{19^2}{9} = 9948$$

d. Bakiye kareler toplamı

$$16565 - 4636 - 9948 = 1981$$

variyans analizi için gerekli tablo aşağıda verilmiştir.

TABLO : 8

Değişikliğin kaynağı	*Kareler toplamı*	*Serb. dere.*	*Tahmini variyans*	*F*
Hane tip. ara.	4636	2	2318.0	3.31
Bölgeler ara.	9948	2	4974.0	10.04
Bakiye	1981	4	495.2	—
Toplam	16565	8	—	—

$\nu_1=2$ ve $\nu_2=4$ serbestlik dereceleri için F tablo değeri $F_{0.05}=6.94$ olduğundan sonuç olarak, bölgeler arasında evet cevapları için görülen farklar manâlı olduğu halde hane tipleri için bu farklar anlamsızdır, diyebiliriz.

Bölgeler itibariyle bir diğer ikili tasnif aile reisinin iş yerindeki durumu göz önünde tutularak yapılmıştır. Neticelerin dökümü aşağıda verilmiştir. Tablo verilerine göre 919 evet'in 540 ı aile reisinin ücretli olarak çalıştığı ailelere *(k)*; 137 si aile reisinin işveren olduğu ailelere *(l)* ve 242 si aile reisinin kendi hesabına çalıştığı ailelere *(m)* aittir.

TABLO : 9

Bölgeler	Aile reisinin işyerindeki durumu (k)	(l)	(m)	Toplam
A	270	62	105	437
B	163	48	76	287
C	107	27	61	195
Toplam	540	137	242	919

Yukarıdaki kıymetlerden hareketle variyans analizinde kullanılacak formüller için gerekli hesaplar yine 100 orijinine ve sıra ve sütunlar arasında karşılıklı tesirin olmadığı faraziyesine göre yapılmış, neticede şu sonuçlar elde edilmiştir.

TABLO : 10

Bölgeler	Aile reisinin işyerindeki durumu (k)	(l)	(m)	T_i	T_i^2	x_{ij}
A	170	—38	5	137	18769	30369
B	63	—52	—24	—13	169	7249
C	7	—73	—39	—105	11025	0889
T_j	240	—163	—58	19	29963	44517
T^2_j	57600	26569	3364	87533		
x^2_{ij}	32918	4977	2122	44517		

Tablo verilerine göre :

a. Sapmaların kareleri toplamı

$$\sum\sum x^2_{ij} - \frac{T^2}{N} = 44517 - \frac{19^2}{9} = 44477$$

b. Aile reisinin işyerindeki durumu için kareler toplamı

$$\sum\left(\frac{T_j^2}{n_{ij}}\right) - \frac{T^2}{N} = \frac{87533}{3} - \frac{19^2}{9} = 29138$$

c. Bölgeler için kareler toplamı

$$\sum\left(\frac{T_i^2}{n_i}\right) - \frac{T^2}{N} = \frac{29963}{3} - \frac{19^2}{9} = 9948$$

d. Bakiye kareler toplamı

$$44477 - 29138 - 9948 = 539$$

TABLO : 11

Değişikliğin kaynağı	*Kareler toplamı*	*Serb. dere.*	*Tahmini variyans*	*F*
Aile rei. işyer. durumu ara.	29138	2	14569	10.8
Bölgeler ara.	9948	2	4974	3.7
Bakiye	5391	4	1348	—
Toplam	44477	8	—	—

Yukarıdaki sonuçlara ve tablo değeri $F_{.0.05}=6.94$ e göre evet'lerin bölgeler arası dağılışı anlamsız iken bu dağılış aile reisinin işyerindeki durumu için manalı çıkmaktadır. Kısaca, (A) marka sabunu kullanan haneler arasında hane reisi ücretli olanların çoğunlukta olduğu söylenebilecektir.

Firma sevk ve idaresi buraya kadar elde olunan sonuçlara bakarak (A) marka sabunun ev kadınları nezdinde işgal ettiği yerin nisbi ehemmiyetini az bulmuş ve kullanımını arttırmak için bir çare olarak sabunun kokusunu değiştirmeyi düşünmüştür. Bu yolda muhtelif koku esansları ile imal edilen ve (A_1) ; (A_2) ; (A_3) ; (A_4) markaları konulan 4 farklı kokuda sabun, bir ön araştırma mahiyetinde olmak üzere, 3 ü A bölgesinden, 3 ü B bölgesinden ve 2 si C bölgesinden olmak şartıyla 1600 soru kâğıdından tesadüfen çekilen 8 ev hanımına birer hafta süreyle kullandırılarak tercihlerini bildirmeleri söylenmiştir. Neticede bu 8 birimden 4 ü (A_3) ü tercih etmiş; 2 si (A_2) yi 2 si (A_4) ü beğenmiş ve (A_1) hiç puan almamıştır. Araştırma servisi bu sonuca dayanarak $[p(A_3)=0.50]$ (A_3) markasının ev kadınları nezdindeki yerinin %50 olduğunu ön görmüş ve neticede bu hüküm aşağıdaki şekilde değerlendirilmiştir.

$$H_0 : p = 0.50$$

$$H_1 : p < 0.50$$

hipotezinin testinde örneğin çok küçük olması nedeniyle normal yaklaşım yerine binom açınımı kullanılmıştır. (A_3) tercihlerinin başarıyı ve $(A_1+A_2+A_4)$ tercihlerinin başarısızlığı temsil ettiği bu problemde, eğer bu 4 marka arasında hiç bir tercih olmasaydı herbiri eşit tercih edilme şansına sahip olacaktı görüşüne göre $(0.75+0.25)^8$ çok terimlisinin açınımı yapılmış ve muhtelif başarı sayıları için ihtimaller hesaplanmıştır.

TABLO : 12

(A_3) ü tercih edenler	*İhtimaller*
1	0.1001
2	0.2670
3	0.3115
4	0.2076
5	0.0865
6	0.0231
7	0.0038
8	0.0004
	1.0000

Bu verilere göre 4 ve daha fazla birimin (A_3) ü tercih etme ihtimali olan son beş ihtimalin toplamı 0.1138, ihtimal eşiği 0.05 den büyük olduğu için (A_3) tercih nisbetinin topluluk için 0.50 olduğu iddiasının kabulü mümkün görülememiştir. Burada hipotez $H_1 : P<0.50$ ifadesini kapsadığından test tek taraflı tatbik edilmiş ve bu sebepten ihtimal olarak 4 ve daha fazla tercih ihtimalleri toplamı alınmıştır.

Bu testin neticesinde firma daha geniş çapta ikinci bir araştırmaya gidilmesini, yeni mamûl (A_3) marka kokunun tercih edilip edilmediğinin tesbiti için daha büyük bir örnekle çalışılmasını arzulamıştır.

Hazırlanan yeni bir çekim şeması üzerinden birbirini takiben sistematik bir şekilde sırasiyle 20; 40; 80; 60 ve 30 cadde ve sokak adı çekilerek herbiri 100; 200; 400; 300 ve 150 birimli 5 küçük örneği kapsayan 1150 birimlik yeni bir nümune elde edilmiştir. Bu örneklere aynı şekilde hazırlanmış sabun çeşitleri yine birer hafta süreyle kullandırılmış ve sonuçta aşağıdaki tasnife varılmıştır.

TABLO 13

	Örnek					
Marka	1	2	3	4	5	Toplam
A_1	15	36	50	69	34	204
A_2	30	55	125	100	45	355
A_3	45	60	200	106	62	473
A_4	10	49	25	25	9	118
Toplam	100	200	400	300	150	1150

Bu problemde bütün örnekler ve ana kütle bölünmesi için tercihler bölünmesinin aynı olduğu hipotezi test edilmekte, görülen farkların manâlı olup olmadığı ki - kare bölünmesinin hesabi değeri ile tablo değerinin mukayesesi suretiyle yapılmaktadır.

$$P(A_1) = \frac{15+36+50+69+34}{1150} = 0.178$$

$$P(A_2) = \frac{30+55+125+100+45}{1150} = 0.308$$

$$P(A_3) = \frac{45+60+200+106+62}{1150} = 0.411$$

$$P(A_4) = \frac{10+49+25+25+9}{1150} = 0.103$$

Bulunan bu nisbetler nümune mevcutları ile ayrı ayrı çarpılarak teorik frekanslar hesaplandı ve değerler tablo şeklinde ifade edildi.

$$e_{11} = 100 \times 0.178 = 17.8$$

$$e_{12} = 200 \times 0.178 = 35.6$$

$$e_{13} = 400 \times 0.178 = 71.2$$

. . . .

. . . .

$$e_{44} = 300 \times 0.103 = 30.9$$

$$e_{45} = 150 \times 0.103 = 15.0$$

TABLO 14

	Örnek					
Marka	1	2	3	4	5	Toplam
A_1	17.8	35.6	71.2	53.4	26.0	204
A_2	30.8	61.6	123.2	92.4	47.0	355
A_3	41.1	82.2	164.4	123.3	62.0	473
A_4	10.3	20.6	41.2	30.9	15.0	118
Toplam	100.6	200.0	400.0	300.0	150.0	1150

$$\chi^2 = \frac{(-2.8)^2}{17.8} + \frac{(0.4)^2}{35.6} + \ldots\ldots + \frac{(-6.0)^2}{15.0} = 82.12$$

12 serbestlik derecesi için $\chi^2_{0.05} = 21.026$ dır. Neticede başlangıç hipotezi olan nisbetlerin eşitliği faraziyesi red edilerek ana kütle için (A_1); (A_2); (A_3) ve (A_4) markaları arasında bir tercih farkı olduğu söylendi. Buna göre firma (A_3) sabununun kokusunu imalâtta esas almayı prensip itibariyle kabul etti; çünkü tercihler bölünmesinde (A_3) nisbeti $P(A_3) = 0.411$ dir.

İstatistik metodların esas gayesi örneklere ait kıymetleri hesaplamak değil, ana kütleye ait değerleri tahmin etmektir. Tesadüfi nümuneler büyüdükçe örnek değerleri limit olarak ana kütle parametrelerine ulaşırlar. Başka bir deyişle, örnek ne kadar büyük olursa bundan hesaplanacak değerler parametrelere o kadar fazla yaklaşırlar. Ahcak tatbikatta bazı hadiselerde örnekler ekseriya küçük olmakta, nümunenin büyüklüğünü arttırmak için genellikle muhtelif küçük tecrübelere ait veriler bir araya getirilmektedir. Fakat böyle bir işleme teşebbüs edebilmek için her şeyden evvel bu küçük tecrübe örneklerinin homojenliklerinden emin olmak lâzımdır. Eğer bu küçük örnekler homojen değillerse, yani aynı ana kütleye ait bulunmuyorlarsa bunlara ait rakkamları bir araya getirerek hesaplanan değerler yığına ait parametreyi karakterize edemezler[21].

O halde biraz evvel hesapladığımız $P(A_3) = 0.411$ değeri acaba herbiri sırasiyle 100; 200; 400; 300 ve 150 birim ihtiva eden beş küçük örneğin içinden çekildiği yığını temsile yeterli midir? Yapılacak bir homojenlik testi sonucunda hesaplanacak ki - kare değeri tablo kıymetinden büyük çıkarsa homojenlik iddiası red edilecek, neticede örneklere ait rakkamları bir araya getirmek suretiyle hesaplanan örnek nisbetinin yığına ait değeri temsil edemeyeceği ileri sürülecektir. Bu yolda ki - karenin «birbirine tâbi olmayan iki ve daha fazla ki - kare'lerin toplamı yine bir ki - kare kıymeti olup bunun serbestlik derecesi bağımsız ki - kare'lerin serbestlik dereceleri toplamıdır» özelliğinden faydalanılmaktadır.

TABLO : 15

Aynı ana kütleden çekilen örnekler	Marka (Fiili Değerler) A_1	A_2	A_3	A_4	Toplam
1	15	30	45	10	100
2	36	55	60	49	200
3	50	125	200	25	400
4	69	100	106	25	300
5	34	45	62	9	150
Toplam	204	355	473	118	1150

21) G.L. Thirkettle «Business Statistics and Statistical Methods» MacDonald Evan Ltd. London 1968 sf. 183.

TABLO : 16

M a r k a

Aynı ana kütleden çekilen örnekler	(Teorik Değerler) A_1	A_2	A_3	A_4	Toplam
1	25	25	25	25	100
2	50	50	50	50	200
3	100	100	100	100	400
4	75	75	75	75	300
5	37.5	37.5	37.5	37.5	150
Toplam	287.5	287.5	287.5	287.5	1150

İkinci tablodaki teorik değerler, koku nev'ileri (marka) arasında bir tercihin yapılmadığı yani ana kütle bölünmesinin 0.25 / 0.25 / 0.25 / 0.25 şeklinde olduğu fikrinde hareketle hesaplanmıştır.

$$\chi^2 = \frac{(15-25)^2}{25} + \frac{(30-25)^2}{25} + \frac{(45-25)^2}{25} + \frac{(10-25)^2}{25} = 30.00$$

$$\chi^2 = \frac{(36-50)^2}{50} + \frac{(55-50)^2}{50} + \frac{(60-50)^2}{50} + \frac{(49-50)^2}{50} = 6.44$$

$$\chi^2 = \frac{(50-100)^2}{100} + \frac{(125-100)^2}{100} + \frac{(200-100)^2}{100} + \frac{(25-100)^2}{100} = 187.50$$

$$\chi^2 = \frac{(69-75)^2}{75} + \frac{(100-75)^2}{75} + \frac{(106-75)^2}{75} + \frac{(25-75)^2}{75} = 54.95$$

$$\chi^3 = \frac{(34-37.5)^2}{37.5} + \frac{(45-37.5)^2}{37.5} + \frac{(62-37.5)^2}{37.5} + \frac{(9-37.5)^2}{37.5} = 39.49$$

+

318.38

		ser. der.
Münferit örnekler χ^2 si toplamları	318.38	15
Toplam olarak 5 örneğin χ^2 si	259.72	2
Nümunelerin birbirinden farklılığından doğan χ^2 kıymeti	58.66	13

Eğer yığın homojen olsaydı bütün örnek ki-kareleri aynı değere yaklaşacak ve neticede fark sıfır olacaktı. Hesabi değerimiz tablo değerinden ($\alpha = 0.05$) büyük olduğu için ($58.66 > 22.362$) homojenlik iddiası red edilmiştir. Bu durumda İstanbul Belediyesi hudutları içindeki bütün hane halkları için (A_3) kokusunun tercih edileceği söylenememektedir. Buna rağmen firma bir süre deneme maksadıyle (A_3) esansını sabun hamuruna koymayı kabul etmiştir. Aradan bir ay geçtikten sonra bu sefer 1600 birimlik eski örnekte evet diyen 919 ev kadınına birer hafta süreyle (A_1); (A_2); (A_3) ve (A_4) markalı sabunlar kullandırılmış ve tercihleri toplanmıştır. Tercihler bölünmesi aşağıdaki gibidir.

TABLO : 17

Marka	*İstanbul Belediye hudutları*	
A_1	119	12.94
A_2	303	33.00
A_3	427	46.44
A_4	70	7.62
Toplam	919	100.00

Bu bölünmenin anlamlı olup olmadığı için ki-kare testi kullanılmış, bu test toplam frekansın şıklar arasında eşit dağıldığı faraziyesini tahkik şeklinde uygulanmıştır.

TABLO : 18

Marka	*İstanbul Belediye hudutları*	
	Fiili	*Teorik*
A_1	119	229.75
A_2	303	229.75
A_3	427	229.75
A_4	70	229.75
Toplam	919	919.00

$$H_0 : f_1 = f_2 = f_3 = f_4 = f_i = 229.75$$

$$H_1 : f_i \neq 229.75$$

hipotezinden hareketle

$$\chi^2 = \frac{(119-229.75)^2}{229.75} + \ldots\ldots + \frac{(70-229.75)^2}{229.75}$$

$$\chi^2 = 53.388 + 23.35 + 169.35 + 111.07 = 357.15$$

bulunur. Üç serbestlik derecesi için $\chi^2_{0.05}=7.815$ olduğuna göre frekansların şıklar arasında eşit bölündüğü faraziyesi kesinlikle kabul edilemeyecektir. Neticede (A_3) ü tercihlerin çoğunlukta olduğu ileri sürülebilir.

Şimdi elimizde iki örnek $(n_1 = 1150 ;\ n_2 = 919)$ ve iki nisbet $(p_1 = 0.411 ;\ p_2 = 0.464)$ bulunmaktadır. (A_3) markasını tercihlerin nisbetinde görülen bu fark neden ileri gelmiştir. Diğer bir ifade ile $p_2-p_1=0.053$ gibi bir fark tesadüfe atfedilebilir mi?

$$H_0 : p_2 = p_1 = p$$

$$H_1 : p_2 \neq p_1$$

Ana kütle nisbetinin tahmini

$$P = \overline{p} = \frac{473+427}{1150+919} = 0.4350$$

$$\sigma_{p1-p2} = \sqrt{P(1-P)\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}$$

$$= \sqrt{(0.43.50)(0.5650)(1/1150+1/919)}$$

$$= \sqrt{(0.2458)\left(\frac{2069}{1.056.850}\right)}$$

$\sigma_{p1-p2} = 0.0219$ bulunur.

$p_1 - p_2 = 0.053$ farkı 2.42 standart hata olduğu ve 1.96 dan büyük çıktığı için $p_2 = p_1$ hipotezi red edilmiştir. Bu şekilde tercihler bölünmesinin (A_3) lehine olduğu kesinleştikten sonra araştırma bu defa bölgeler itibariyle ele alınmış, (A) marka sabun kullanma nisbetinin en düşük olduğu bölge olan B de kokunun değiştirilmesinin talebi ne şekilde etkilediği incelenmiştir.

TABLO : 19

Marka	*B*	*Bölgesi*	*İstanbul*	
A_1	64	22.30	119	12.94
A_2	85	29.62	303	33.00
A_3	88	30.66	427	46.44
A_4	50	17.42	70	7.62
Toplam	287	100.00	919	100.00

Veriler gözden geçirilirse iki dağılma arasında bariz bir farklılık olduğu süratle belirmektedir. Gerçekten (A_3) ü tercih edenler nisbeti İstanbul için 0.4644 iken bu değer B bölgesi için 0.3066 dır. Ayrıca B bölgesinde tercihler arasında büyük sapmalar yoktur.

$p_1 = 0.4644$ $p_B = 0.3066$ arasında görülen 0.1578 gibi bir fark tesadüfi olabilir mi? Yoksa B bölgesi tercihi tüm yığma nazaran esaslı bir aykırılığa mı sahiptir?

Bu konuda :

a. p_2-p_1 testi uygulanabilir. $p_2=p_B=0.3066$ $p_1=p_I=0.4644$ dür. İkinci nisbet tüm yığma ait bir değer olduğuna göre bu test $p-P$ farkının tahkikine dönüşmekte ve neticede kullanılacak standart hata

$$\sigma_p = \sqrt{\frac{P(1-P)}{n}} = \sqrt{\frac{(0.4644)\,(0.5356)}{287}}$$

den hesaplanmaktadır.

$$z = \frac{0.1578}{0.499/17} = 5.37 \quad \text{elde edilmekte}$$

bu değer ($\alpha = 0.05$) 1.96 dan büyük çıktığı için 0.1578 farkının tesadüfen meydana gelmediği, anlamlı bir fark olduğu söylenebilmektedir. Bu sonuca göre B bölgesindeki tercihler bölünmesi İstanbul Belediye hudutları içindeki tercihler bölünmesinden ayrılmakta, B bölgesindeki aileler için kokunun farklılaştırılması hiç bir önem kazanmamaktadır.

b. Yukarıdaki test her marka için ayrı ayrı uygulanabilirdi. Bu takdirde test her marka için aynı sonucu veriyorsa «anlamlı veya anlamsız» iki dağılma arasındaki farkın anlamlı veya anlamsız olduğu söylenebilir.

Gerçekten her marka için z değerleri hesalpanırsa

$$\text{Marka } (A_1) \qquad z_1 = \frac{0.2230-0.1294}{\sqrt{\frac{(0.1294)\,(0.8706)}{287}}}$$

$$= \frac{0.0936}{0.336/17} = 4.73$$

$$\text{Marka } (A_2) \qquad z_2 = \frac{0.2962-0.3300}{\sqrt{\frac{(0.33)\,(0.67)}{287}}}$$

$$= \frac{0.0338}{0.47/17} = 1.22$$

Marka (A_4) $$z_4 = \frac{0.1742 - 0.0762}{\sqrt{\frac{(0.0762)(0.9238)}{287}}}$$

$$= \frac{0.098}{0.266/17} = 6.28$$

kıymetleri bulunmaktadır. Bu durumda test her marka için aynı sonucu vermediğinden iki dağılma arasındaki farkın anlamlı veya anlamsız olduğu söylenememiştir. Bunun üzerine

c. Ki-kare testinin tatbikine girişilmiştir. Yapılacak tahlilde kullanılacak verilerin dökümü aşağıdadır.

TABLO : 20

Fiili Değerler

Marka	*(A+C) Bölgesi*	*B Bölgesi*	*İstanbul*
A_1	55	64	119
A_2	218	85	303
A_3	339	88	427
A_4	20	50	70
Toplam	632	287	919

Teorik değerler, İstanbul için nisbetler bölünmesinin (A+C) ve B bölgesinde aynen geçerli olduğu görüşünden hareketle hesaplanmıştır.

TABLO : 21

Teorik Değerler

Marka	*(A+C) Bölgesi*	*B Bölgesi*	*İstanbul*
A_1	81.8	37.2	119.0
A_2	208.4	94.6	303.0
A_3	293.7	133.3	427.0
A_4	48.1	21.9	70.0
Toplam	632.0	287.0	919.0

$$\chi^2 = \frac{(55-81.8)^2}{81.8} + \frac{(64-37.2)^2}{37.2} + \frac{(218-208.4)^2}{208.4} + \ldots$$

$$\ldots + \frac{(50-21.9)^2}{21.9}$$

$$\chi^2 = 8.78 + 19.31 + 0.44 + 0.97 + 6.98 + 15.39 + 16.40 + 36.05 = 104.32$$

Bu netice bize (A_3) tercihlerinin bölgeler itibariyle eşdeğer olmadığı sonucunu vermiş, (A) marka sabunu kullanma nisbetinin yüksek olduğu A ve C bölgeleri için (A_3) tercihleri 0.50 nin üzerinde olduğundan (0.536) firma sevk ve idaresi (A_3) kokusunu imalâtta esas almayı prensip itibariyle kabul etmiştir.

IV

Günümüz yöneticisinin, işletme idaresi tekniğindeki gelişmelere parelel olarak, yüklendiği sorumluluklar her geçen gün biraz daha artmakta, maksimum sonuçların minimum bir faaliyetle sağlanması yolunda sarfedilen çabalar işletme içi ve dışı güvenilir verilere istinaden yapılan istatistik analizlerle değer kazanmaktadır. En verimli hedeflerin tertibinden başlayıp bu hedeflere erişinceye kadar geçen süre içinde ve bilhassa tahmin, plânlama ve kontrol safhalarında kullanılan istatistik metodları işletme istatistiği başlığı altında toplanır.

Faaliyet gayesi ve çalışma sahası ne olursa olsun günümüz işletmelerinin uzun vadeli ve kârlı bir şekilde gittikçe büyüyerek yaşayabilmeleri, ekonomik ve teknik gelişmelerden doğacak değişmelere karşı her zaman hazırlıklı bulunmaları ile mümkün olabilir. İşletmeleri ve tabiatıyle sevk ve idareyi, muhtemel değişikliklerin getireceği tehlikelere karşı daima hazır bulunduracak ve hatta onların kendi menfaatlarına uygun olacak şekilde değişiklikler hazırlamasına imkân verecek en önemli yardımcıları pazarlama araştırmalarıdır.

Sevk ve idare için temel fonksiyonu «karar almada yardımcı olma» şeklinde özetlenen bir pazarlama araştırması daha geniş tarifini «mal ve hizmetlerin üreticiden tüketiciye intikali veya satışı ile ilgili problemlere ait bütün verilerin toplanması, kaydedilmesi ve analizi» ifadesinde bulunur.

«Biri evet diğeri hayır olmak üzere sadece iki muhtemel sonuç veren ve aynı şartlar altında, her seferinde elde edilen neticeler yekdiğerinden bağımsız olacak şekilde *(n)* defa tekrarlanan bir hadiseye ait verilerin yorumlanmasında kullanılan örnekleme prensiplerinden» istifade edilecek şekilde hazırlanan pazarlama araştırması çalışmamızın esasını teşkil etmiştir. Tüketici piyasasının işletmenin (A) marka banyo sabunu açısından incelenmesi; pazarları ele geçirmek ve pazar payını arttırmak için neler yapılması gerektiği; yeni bir mamûl yapıldıktan sonra bu mamûlün pazarın ihtiyacını karşılayıp karşılamadığını görmek düşüncesi; mevcutların tüketiciler tarafından âzami kabul görecek şekilde değiştirilmesi gayesi; satış ve reklâm çabalarının yoğunlaştırılacağı yerlerin seçimi ile yapılan reklâmın amaca ulaşıp ulaşmadığını bilmek ve muhtemel alıcıları ne ölçüde etkilediğini ölçmek amacı veya genel olarak yöneticinin faaliyette bulunduğu belirsizlik alanının daraltılması, karar vermede hata ihtimalinin minimuma indirilmeye çalışılması prensipleri bu araştırmanın itici unsurları olmuştur.

KAYNAKLAR

G. Calot (Terc. E. Güntekin) «Örnekleme Metodlarının Meskenle İlgili Problemlere Uygulanması» Ankara 1969.

H. Arkın - R.R. Colton (Terc. S. Kendir) «İstatistik Metodlar» Ankara 1968.

C.E. Weatherburn «Mathematical Statistics» Cambridge University Press 1957.

F. Conway «Sampling An Introduction for Social Scientist» London 1967.

G.W. Snedecor «Statistical Methods» The Icwa State College Press 1950.

J.G. Peatman «Introduction to Applied Statistics» London 1963.

F.E. Croxton - D.J. Cowden «Applied General Statistics» London 1955.

M.R. Spiegel «Theory and Problems of Statistics» Schaum 1961.

R.G. Allen «Statistics for Economists» Hutchinson Ltd. London 1966.

A.M. Turtle «Elementary Business and Economic Statistics» MacGraw Hill 1957 London.

R. Loveday «A Second Course in Statistics» Combridge University Press 1964.

W.G. Cochran «Comparison of Percentages in Matched Samples» Biometrica 37-1950.

R. Goodman «Statistics» The English Universities Press London 1960.

G.L. Thirkettle «Business Statistics and Statisticai Methods» MacDonald Evan Ltd. London 1968.