# Kİ-KARE VE KOLMOGOROV-SMİRNOV UYGUNLUK TESTLERİ VE BUNLARIN KARŞILAŞTIRILMASI

Mustafa AYTAÇ*

Yapılan araştırmanın sonucunda uygulanan istatistik testin kuvvetli olabilmesi için, bu testin en güçlü ya da en kapsamlı varsayımları içermesi gerekir. Bu varsayımlar geçerli olduğu zaman, istatistik testin gücü, alabildiği en yüksek değere çıkar. Diğer bir deyimle açıklamak istersek, bu en güçlü ve en kapsamlı varsayımlar geçerli olduğu zaman, bu testlerin diğer testlere nazaran, sıfır hipotezi yanlış olduğunda reddetme olasılığı, en yüksek olanıdır.

Parametrik istatistik testleri genellikle ortalamaların ve varyansların testlerinde kullanıldığı gibi bir, iki ve ikiden çok yığına da aynı güç etkinliğinde uygulanabilir. Çünkü parametrik testler, kullanışlarını belirleyen çok güçlü varsayımlara sahiptirler Buna en güzel örnek t- testi veya F- testidir.

Parametrik istatistik testleri uygularken kabul edilen varsayımlar şunlardır:

a) Gözlemler birbirlerinden bağımsız olmalıdır. Bir diğer ifade ile yığından alınıp gözlenen birimin seçilme şansı ve bu birimin aldığı puan, yığındaki diğer birimlerin seçilme şansını ve puanını etkilememelidir.

b) Parametrik istatistik testleri belli durumlarda — Normal dağılım gibi — sürekli bir dağılımın varlığını kabul ederler.

c) Yığının şekli konusunda bir şüphe varsa, örnek sayısı merkezi limit teoreminden yararlanabilmek amacıyla otuzdan fazla seçilmelidir.

d) Birimlerin aldığı puanlar üzerinde aritmetik işlemlerin yapılabilmesi için incelenen değişkenler en azından aralıklı ölçek düzeyinde ölçülmüş olmalıdırlar.

e) Yığınlar aynı varyans veya özel durumlarda bilinen bir varyans oranına sahip olmalıdırlar.

Parametrik testleri, varyans analizinde kullanmak istiyorsak bu beş varsayıma, aşağıdaki varsayımı da eklememiz gerekmektedir.

f) Normal veya eşit varyansa sahip yığınların ortalamaları, sütun veya sıralardan doğan etkilerin doğrusal bileşkeleri olmalıdır. Yani etkiler toplanabilmelidir.

Son iki varsayımı bazı özel durumlarda öne sürülebildiğinden bir kenara bırakırsak, ilk dört varsayım önceden kabul edilen varsayımlardır. Bu varsayımlardan herhangi bir tanesi uygun değilse uygulanan parametrik testler yanıltıcı sonuçlar verebilir. Bu nedenden dolayı küçük hacimli örneklerle çalışırken her türlü bölünüm için geçerli olabilen test yöntemlerine gerek vardır.

---

(*) Bursa Üniversitesi İktisadi ve Sosyal Bilimler Fakültesi Yardımcı Doçenti.

Parametrik olmayan istatistik testleri, klasik parametrik istatistik testlerinin bir seçeneği olarak ortaya çıkmıştır [1]. Parametrik olmayan testler II. Dünya Savaşından sonra çok gelişmiş ve hemen hemen birçok alanda uygulanma olanağı bulabilmiştir.

Bir çok istatistikçi parametrik olmayan (Nonparametric) ve serbest dağılımlı (Distribution-Free) testleri aynı ad altında kullanmaktadır. Fakat aslında ikisinin arasında küçük bir fark vardır [2].

Parametrik olmayan testler, parametrelerin özel değerleri üzerine hiç bir varsayım yapamadığımız durumlarda uygulanabilir. Biraz sonra inceleyeceğimiz KOLMOGOROV-SMİRNOV uygunluk testi buna bir örnektir. Serbest dağılımlı testler ise, yığının biçimi ve şekli hakkında bir varsayım olmadığı durumlarda kullanılır. Ki-Kare uygunluk testi de bunun için bir örnektir.

Ayrıca şunu da üzerinde önemle durarak belirtmemiz gerekir ki, parametrik testlerin uygulanılacağı durumlarda parametrik olmayan test yöntemlerine başvurulmamalıdır. Çünkü parametrik test yöntemleri, parametrik olmayan test yöntemlerine göre daha kesin sonuçlar verir. Başka bir ifade ile herhangi bir $\alpha$ anlamlılık seviyesinde bir testin sahip olduğu II. tip hata $\beta$; parametrik olmayan istatistik testlerde, parametrik istatistik testlerinden daha büyük çıkacaktır.

1- Tek Örneklemli Durumlarda Uygunluk Testleri:

Araştırmacıların, uygulamalarda sık sık karşılaştıkları sorunlardan birisi; ele alınan n birimlik bir örneğin hangi dağılıma sahip bir yığından geldiğinin bulunmasıdır. Bu türdeki sorunların çözümünde yararlanılan testler, uygunluk testleridir.

Uygunluk testindeki sıfır hipotezi, yığının olasılık fonksiyonu üzerine kurulmuş bulunan hipotezdir. Almaşık hipotez ise sıfır hipotezinin geçerli olmadığı hipotezdir.

Çalışmamızda parametrik olmayan istatistik testlerden Ki-Kare ve Kolmogorov-Smirnov uygunluk testlerini tek örneklemli durumlarda incelemeye ve bu iki testin karşılaştırmasını yapmaya çalışacağız. Her iki test de hipotez edilen teorik dağılım ve gözlenen birimlerin dağılımı arasındaki karşılaştırmaya dayandırılmıştır. Bununla beraber iki test de karşılaştırmayı yaparken birbirlerinden ayrı olan yöntemleri kullanırlar. Bu iki testi ele almamızın nedeni, bu iki testin daha önemli ve diğer uygunluk testlerine temel olmalarından dolayıdır.

Tek örneklemli durumlarda kullanılan başka uygunluk testleri de vardır. Bunlar David Uygunluk Testi [3], Lilliefors Uygunluk Testi ve Cramer von Mises Uygunluk Testleridir [4].

David Uygunluk Testi, Kolmogorov-Smirnov Uygunluk Testinin bir seçeneği olarak düşünülebilir. Lilliefors ve Cramer von Mises Uygunluk Testleri ise Kolmogo-

---

1 Parametrik olmayan istatistik yöntemi ilk olarak işaret (sign) testi ile uygulayan 1710 yılında Arbuthnot olmuştur. Lehmann, E.L.: Nonparametrics, Holden-Day, 1975, S. VII.

2 Freud, F.E. and Williams, F.J.: Elementary Business Statistics, Prentice-Hall, 1969, S: 352-354.

3 James V. Bradley, Distribution-Free Statistical Tests, Prentice-Hall Inc, New Jersey 1968, s. 304-310.

4 William Jay Conover, Practical Nonparametric Statistics, John Wiley, New York, 1959, s. 302-309.

rov-Smirnov uygunluk testinin varsayımlarına dayanmakla birlikte bazı küçük farklılıklar gösterirler. Bu farklılıklar ise kısaca şunlardır: Lilliefors Uygunluk Testinde, test istatistiği gözlem değerlerinden doğrudan doğruya hesaplanmaz. Bunun yerine,

$$Z_i = \frac{x_i - \bar{x}}{s} \qquad i = 1, 2 \ldots\ldots, n$$

dönüşümü yapılarak $Z_i$ değerlerinden bulunur. Cramer von Mises Uygunluk Testinde ise, gözlem değerleri sıralayıcı istatistikler olarak sıralandıktan sonra ele alınırlar.

### 1.1. Ki-Kare Dağılımı ve Ki-Kare Uygunluk Testi:

En eski ve en iyi uygunluk testi olarak bilinen Ki-Kare uygunluk testi, ilk olarak 1900 yılında Pearson tarafından ele alınmıştır [5].

Ki-Kare Uygunluk Testinin nasıl ve ne şekilde kullanıldığına geçmeden önce, ispatlarını vermeden kısaca nasıl elde edildiğine değinelim.

X'in Gamma fonksiyonu $\Gamma(x)$ olarak gösterilir ve şöyle ifade edilir.

$$\Gamma(x) = \int_0^{\infty} e^{-t}\, t^{x-1}\, dt \qquad x > 0$$

Gamma $\Gamma(x)$'in olasılık fonksiyonu ise şöyledir:

$$f(x;\alpha, n) = \begin{cases} \dfrac{\alpha^n e^{-\alpha x} x^{n-1}}{\Gamma(n)} & ; x > 0 \\ 0 & ; x \leqslant 0 \end{cases}$$

$\alpha$ ve $n$, gamma olasılık fonksiyonunun parametreleri olup, $\alpha > 0$ ve $n > 0$ koşulları da sağlanmalıdır.

Açıklıkla gösterilebilir ki [6],

$$f(x) \geqslant 0 \quad \text{ve} \quad \int_{-\infty}^{\infty} f(x;\alpha, n)\, dx = 1 \text{ dir.}$$

Gamma olasılık fonksiyonunda iki parametrenin özel değerler alması Ki-Kare olasılık fonksiyonunu ortaya çıkarır. $\alpha = \frac{1}{2}$ ve $n = \frac{m}{2}$ ($m > 0$ ve tamsayı) olduğu zaman $f(x;\alpha, n)$ aşağıdaki gibidir.

$$f\left(x;\frac{1}{2}, \frac{m}{2}\right) = \begin{cases} \dfrac{e^{-1/2 x}\, x^{(m/2-1)}}{2^{m/2}\, \Gamma(m/2)} & ; \quad x > 0 \\ 0 & ; \quad x \leqslant 0 \end{cases}$$

$f(x; 1/2, m/2)$ olasılık fonksiyonu, $m$ serbestlik derecesi ve sürekli bir $x$ değişkenine sahip Ki-Kare dağılımıdır [7]. $\chi^2_{(m)}$ ile gösterilen Ki-Kare olasılık fonksiyonunda $(m)$ aynı zamanda fonksiyonunun parametresidir.

Genel olarak Ki-Kare birikimli dağılım fonksiyonu olan

$$F(X;\alpha, n) = \int_0^{x} t(x;\alpha, n)\, dt$$

olasılığı hesaplanmasındaki güçlükten dolayı hazırlanmış Ki-Kare tabloları kullanılarak bulunabilir.

---

5 Conover, A.g.e., s. 186.
6 Bernard Harris, Theory of Probability, Addison-Wesley, 1966, s. 65-69.
7 Robert V, Hoggy-Allen T. Craig: Introduction to Mathematical Statistics, Mac Millan, Hong-Kong, Third Edition, 1972, s. 99-103.

Ki-Kare testi,uygunluk testinden başka, bağımsızlık testi, türdeşlik testi, belirli bir hipoteze uygunluk testi ile korelasyonun doğrusallığının araştırılması testinde de kullanılır [8].

Ki-Kare uygunluk testi, ele aldığı örneklerin özelliklerine göre sınıflandırma yoluna gidilerek yapılır. Bu sınıflandırma yapılırken, şüphesiz ki bir özelliğin diğerlerine oranla daha sık ortaya çıkacağı önceden düşünülebilir. Ayrıca sınıflama sayısı ikiye eşit veya daha fazla olabilir. Bu test her sınıfa düşen nesne veya cevapların gözlenen sayısı ile, sıfır hipotezine bağlı olarak beklenen sayı arasında "anlamlı" bir farkın olup olmadığını test etmek için kullanılır.

Cramer ve Birnbaum, "Geliştirilmiş Minimum Ki-Kare yöntemiyle (Modified minimum Chi-Square Method)", Ki-Kare uygunluk testinin kullanılmasında akılcı çözümler getirmişlerdir. Bu yöntemde her sınıflama birbirine eşit uzunlukta ele alınıp, bu sınıflamalardaki veriler ile parametrelerin tahminleri hesaplanarak bir olasılık dağılım tablosu elde edilmektedir. Elde edilen bu olasılıklara göre hipotez edilen gözlem değerleri bir sınıflamaya konulmaktadır. Sonuçta yine gözlem değerlerinin dağılımı ile hipotez edilen dağılım arasında bir karşılaştırma yapılmaktadır [9].

### 1.1.1. Ki-Kare Uygunluk Testinin Varsayımları:

Ki-Kare uygunluk testini kullanmadan önce aşağıdaki varsayımları kabul ederiz.

a) Gözlem değerleri bağımsız olup, herhangi bir yığından $n$ sayıda tesadüfi olarak ele alınmışlardır.

b) Verilerin ölçme düzeyi sınıflandırıcı olabilir.

Sınıflandırıcı ölçek de, gözlenen nesneler belirli özelliklere göre sınıflanır. Her sınıf bir sayı veya başka bir işaretle gösterilir. Bir sınıfın numarasının diğerinden büyük veya küçük olması bir anlam taşımaz, yalnızca iki sınıfın birbirinden ayrı olduğunu belirtir. Bu nedenden dolayı bir sınıftaki nesneler elden geldiğince aynı yapılmaya çalışılır. Taşıtlara, illere göre numara verilmesi gibi.

c) Gözlem değerleri bütün sınıflandırma olanaklarını içeren ve birbiri ile çakışmayan $k$ tane sınıfta sınıflandırılabilir.

### 1.1.2. Ki-Kare Uygunluk Testinin Hipotezleri ve Test İstatistiği:

$F(x)$, X tesadüfi değişkeninin gerçek fakat bilinmeyen dağılım fonksiyonu olsun. $F^*(x)$'de X'in tam olarak belirlenmiş hipotez edilen dağılım fonksiyonu olsun. O zaman hipotezlerimiz,

$$H_0 = F(x) = F^*(x) \qquad H_1 = F(x) \neq F^*(x)$$

veya kelimelerle ifade etmek istersek,

$H_0$: Gözlenen tesadüfi değişkenlerin dağılım fonksiyonu $F^*(x)$'dir.

$H_1$: Gözlenen tesadüfi değişkenlerin dağılım fonkisyonu $F^*(x)$'den farklıdır.

Hemen görüleceği gibi Ki-Kare uygunluk testindeki hipotezimiz yalnızca çift taraflıdır. Tek taraflı testler bizi ilgilendirmez. Başka bir ifade ile almaşık hipotezimiz, hipotez edilen dağılımın gerçek dağılımdan hangi yönde farklılık gösterdiği problemi ile ilgilenmez ve böyle bir soruya açıklık getirmez.

---

8 M. Kemal Yoğurtçugil: Ki-Kare üzerine bir deneme-İst. Üniversitesi İktisat Fak. Yayınları, İstanbul, 1978, s. 11-14.

9 Conover, A.g.e. s. 191-194.

X'in dağılım fonksiyonu $F^*(x)$ dir. Varsayım altında $i$'inci sınıftaki X tesadüfi gözleminin olasılığını $P_i$ olarak gösterelim. O zaman $B_i$'yi şu şekilde tarif edebiliriz.

$$B_i = P_i \cdot n_i \qquad i = 1, 2 \ldots, k$$

Diğer bir anlatımla $B_i$, $H_0$ hipotezine göre $i$'inci sınıfta beklenen birim sayısıdır. O zaman örnekten alınan birimlerin Ki-Kare hesabındaki test istatistiğimiz:

$$\chi^2 = \sum_{i=1}^{k} \frac{(G_i - B_i)^2}{B_i}$$

veya

$$\chi^2 = \sum_{i=1}^{k} \frac{G_i^2}{B_i^2} - n \qquad \text{dir.}$$

Formülde:

$G$: $i$'inci sınıftaki gözlenen birim sayısı,

$B$: $H_0$ hipotezine göre, $i$'inci sınıfta beklenen birim sayısıdır.

$k$: Gözlemlerin toplandığı sınıflama sayısıdır.

Daha öncede açıkladığımız gibi $\chi^2$ tablosu, sürekli teorik dağılımından hesaplandığı halde, formülde kesikli durumlar için hesaplama yapılmaktadır. Bu sorunu süreklilik için Yates Düzeltmesi (Yates' Correction) ile çözmekteyiz [10].

Yates'in

$$\chi^2 = \sum_{i=1}^{k} \frac{(|G_i - B_i| - \frac{1}{2})^2}{B_i}$$

formülü kullanıldığı zaman, $\chi^2$ yaklaşımı düzeltilebiliniyor. Bu düzeltme yalnızca Ki-Kare dağılımı bir serbestlik derecesine sahip olduğu zaman kullanılabilir. Eğer serbestlik derecesi birden fazla olursa kullanılmaz. Aynı zamanda değişik Ki-Kare değerleri birleştirildiği zaman da dikkate alınmayabilir.

Formülden de görüldüğü gibi $B_i$ değerleri büyük olduğu zaman 1/2 düzeltmesi az etki eder. Fakat $B_i$ küçükse o zaman önemli bir düzeltmedir. Yates düzeltmesi $(G_i - B_i) < 1/2$ olduğu zamanlarda dikkate alınmayabilir. Ki-Kare değeri büyük olursa gözlenen sıklıkların $H$ hipotezinin dayandığı yığından gelmiş olması daha az olasıdır ve Ki-Kare değerini etkileyen temel faktör $(G_i - B_i)$ farkıdır.

Ki-Kare uygunluk testi değişik seçeneklerde daha iyi yararlanılabilmesi için, belirli bir olasılığa uygunluk testlerine de dönüştürülebilir [11].

### 1.1.3. Ki-Kare Uygunluk Testinde Karar Alma:

Test istatistiğimizdeki $k$ harfi, sınıflandırmada ki sınıf sayısını ve Ki-Kare örnek dağılımının serbestlik derecesini (s.d) belirler.

Ki-Kare tablo kritik değerinin bulunmasında önemli rol oynadığı için s.d. üzerinde biraz duralım.

Ki-Kare dağılımının s.d'si $(k - 1)$ dir. Bunun nedeni $k$ sınıfdan, $(k - 1)$ sınıfın birim sayısını bildiğimiz zaman $k$'inci sınıftaki birim sayısını kolaylıkla belirleyebile-

---

10 Taro, Yamane, Statistics and Introductory Analysis, Harper International Edition, Singapore, Third Edition, 1973, s. 766-777.

11 Bakınız: Yoğurtçugil, A.g.e., s. 27-61.

ceğimiz içindir. Örneğin 10 öğrenciyi iktisat, işletme, sosyal siyaset ve Koop. ile siyasal bilim bölümlerini seçenler diye $k = 4$ sınıfa ayırsak; ve İktisadı 3 kişi $(A_1)$, İşletmeyi 2 kişi $(A_2)$ ve sosyal siyaset ve Koop. 3 kişi $(A_3)$ seçmişse, siyasal bilimleri seçen öğrenci sayısı $A_4 = [n - (A_1 + A_2 + A_3)]$'dür. Bundan dolayı bu örnek için s.d'si üçtür. Diğer bir ifade ile, s.d. $k - 1$ eşitliğinden hareket edersek; s.d = 4-1 = 3 dür.

Ki-Kare uygunluk testini uygulayacağımız zaman, $P_i$ olasılıklarını hesaplamadan önce çoğunlukla hipotez edilen yığının belirli parametrik tahminlerini örnekten hesaplamamız gerekebilir. Bu durumlarda test istatistiğimizi daha önceki yöntemle buluruz. Fakat s.d'si farklı olacağı için şu şekilde hesaplarız: Tahmin edeceğimiz her parametre için önceki s.d.'den bir çıkarırız. $L$ harfi örnekten tahmin edeceğimiz parametre sayılarını gösteriyorsa bu durumda s.d. miz $(k - 1 - L)$ dir.

Bu testde karara varmak için, test istatiğimizde bulunan değer ile $k - 1$ (veya $k - 1 - L$) s.d'si ve $\alpha$ önem seviyesinde Ki-Kare tablolarından bulunan değer karşılaştırılır. Hesaplanan $\chi^2$ değeri, tablodan bulunan Ki-Kare değerine eşit veya ondan büyükse, $H$ hipotezinin $\alpha$ önem seviyesinde reddederiz.

### 1.1.4. Ki-Kare Uygunluk Testinde Beklenen Frekansların ($B_i$) Küçük veya Serbestlik Derecesi Büyük Olduğu Zaman:

Ki-Kare uygunluk testinin uygulandığı alanlarda çok sık rastlanılan yanlış kullanma hatalarından birisi de, beklenen frekansların küçük olmasıdır. Süreksiz bölünme olan binomun, normal bölünmeye yaklaşımı ile sağlanan yararlar, aynı neden ile Ki-Kare bölünmesinin kullanılması ile de sağlanmaktadır. Ancak kullanım beklenen birimlerin 10'dan büyük olması ile sınırlanmıştır [12].

S.d'sinin bire yakın olduğu durumlarda, beklenen birimler 5'den küçük olursa o sınıfı bir önceki veya sonraki sınıfla birleştirerek sınıflama sayısını bir eksiltme yoluna gidilmelidir.

Beklenen birimlerin 5'den küçük olması koşulu çeşitli yazarlarca 10'dan ve hatta Kendall'a göre 20'den az olmama şeklinde daha katı olarak uygulamaya alındığını görmekteyiz [13].

Beklenen birimlerin — özellikle s.d'si 5'den küçükse — 5'den büyük olabilmesi için gruplama olanakları aranarak büyük yapılmaya çalışılmalıdır. Gruplama yapılamıyorsa bu durumda hipergeometrik dağılıma dayanan ve Fisher Kesinki — Kare Analizi adı verilen bir test uygulanır. Bu analiz, hesaplanan faktöriyeller veya büyük sayılarda, bunların logaritmaları üzerinden yapılmakta, en küçük birim sıfıra ininceye kadar marjinal toplamlar değişmemek koşulu ile yeni tablolar hazırlanmaktadır [14].

Ki-Kare dağılımının s.d.'si 30'dan büyük olduğu durumlarda, normal yaklaşım kullanılır. Çünkü tablolaştırılmış Ki-Kare değerleri en fazla s.d.'si 30 olmak üzere yapılmıştır. Bu durumda aşağıdaki normal yaklaşımlardan birisini ve normal tabloları kullanarak karara varılır [15].

---

12 Yoğurtçugil, A.g.e., s. 23.
13 Aynı eser, s. 42.
14 Aynı eser, s. 23.
15 Harris, a.g.e., s. 22.

Fisher yaklaşımı: X, $n$ s.d.'inde bir Ki-Kare dağılımına sahipse, o zaman

$$Z = \sqrt{2x} - \sqrt{2n-1}$$

değişkeni $\mu = 0$ ve $\sigma^2 = 1$ olmak üzere yaklaşık normal dağılır.

Wilson-Hilferty yaklaşımı: X, $n$ s.d'nde bir Ki-Kare dağılımına sahipse, o zaman

$$Z = \sqrt[3]{\left(\frac{x}{n}\right)} - \frac{1-(2/9n)}{\sqrt{(2/9n)}}$$

değişkeni $\mu = 0$ ve $\sigma^2 = 1$ olmak üzere yaklaşık normal dağılır.

Yukarda verilen iki yaklaşımdan en iyisini Wilson-Hilferty yaklaşımı vermesine rağmen, hesaplanmasındaki zorluktan ve bıktırıcı olmasından dolayı, uygulamada çoğunlukla Fisher yaklaşımı kullanılır.

S.d.'si 30'dan büyük olduğu durumlarda Fisher yaklaşımındaki $Z$ değeri hesaplanarak, standart normal tablodan bulunan değer ile karşılaştırılıp bir karara varılır.

### 1.1.5. Ki-Kare Uygunluk Testinin Etkinliği:

Sınıflayıcı ölçme kullanıldığında veya verilen aynı sınıflardaki birimlerden oluştuğunda, Ki-Kare testinin güç etkinliği kavramı anlamsızdır. Çünkü böyle durumlarda uygulanabilecek herhangi bir test yoktur. Eğer veriler parametrik bir testin uygun düşeceği şekilde ise, o zaman Ki-Kare testinin kullanılması bilgilerin harcanması demektir.

S.d.'si birden büyükse Ki-Kare testi sıralanma etkilerine duyarlı değildir ve bundan dolayı da, eğer hipotez sıralanmayı gözönüne alıyorsa, Ki-Kare testinin en iyi test olmadığına dikkat etmek gereklidir.

## 1.2. Kolmogorov-Smirnov Tek Örneklem Uygunluk Testi:

K-S uygunluk testi ilk olarak 1933 yılında bir Rus matematikçisi olan Kolmogorov tarafından ortaya atılmıştır [16]. Kolmogorov $X_1, X_2, \ldots X_n$ tesadüfi değişkenlerinin bilinen sürekli bir dağılım fonksiyonu $F(X)$'in, uygunluk problemleri için uygun test olan

$$D_n = n^{1/2} \sup_{-\infty < x < \infty} |F_n(x) - F(x)|$$

istatistiğinin limit dağılımını bulmuştur.

1939 yılında başka bir Rus matematikçisi Smirnov iki örneklemli durumlar için bunu geliştirmiştir. 1945 yılında da Massey

$$K^{\pm} = \left(\frac{mn}{m+n}\right)^{1/2} \max_{-\infty < x < \infty} |F_{2,n}(x) - F_{1,m}(x)|$$

K-S test istatistiğini $m \leqslant 40$, $n \leqslant 40$ ve yine $m \neq n$ $(m \leqslant 10, n \leqslant 10)$ örnek büyüklükleri için geliştirmiştir. Seçilen büyük örnekler için bunun asimtotik dağılımı Smirnov tarafından 1939 ve 1948'de ayrıca Kolmogorov tarafından da 1954 yılında geliştirilmiştir. Bu test istatistiğinin tabloları ise 1960 yılında Birnbaum-Hall tarafından $m \leqslant 40$, $n \leqslant 40$ için bir araya getirilmiştir.

---

16 Bundan sonra Kolmogorov-Smirnov'un baş harflerini (K-S) kullanacağız.

K-S tek örneklem uygunluk testini uygulayacağımız zaman iki birikimli dağılım fonksiyonu üzerine dikkatlerimizi toplarız: bunlar hipotez edilen teorik birikimli dağılım ile gözlenen birikimli dağılımdır. Birikimli dağılım fonksiyonu $F(X)$, verilen belli bir $X$ değeri için rastgele değişken $X$ değerinin $X$'e eşit veya küçük olması olasılığıdır. Yani $F(X) = P\,(X \leqslant x)$.

Dağılım fonksiyonu bilinmeyen $F(X)$'den rastgele bir örnek çekelim. $X$'in her değeri için $F(X)$ ile hipotez edilen teorik birikimli dağılım fonksiyonu $F_0(x)$ arasında bir fark olup olmadığını araştırırız. Eğer $F(x) = F_0(x)$ ise, $F_0(x)$ ile gözlenen dağılım fonksiyonu (empirical distribution function) olan $S_n(x)$ arasında yakın bir ilişki olmasını bekleriz. K-S tek örneklem uygunluk testinin temeli $F_0(x)$ ile $S_n(x)$ arasındaki ilişkinin olup olmadığını belirleyebilmek için $F(x) = F_0(x)$ sıfır hipotezinin "anlamlı" olup olmadığına bakar.

Gözlenen dağılım fonksiyonu $S_n(x)$'i aşağıdaki gibi tanımlarız [17]. Gözlenen dağılım fonksiyonu, $x$'in bütün değerleri için belirlenmiş $n$ hacimli bir rastgele örneğin dağılım fonksiyonudur.

Böylece $S_n(x)$ örneğin sıralayıcı istatistikleri olan (order statistics) atlama noktalarında (Jump points), $1/n$ çarpanı ile gittikçe çoğalan bir basamak fonksiyonudur. Tesadüfi bir örneğin sıralayıcı istatistiklerini $X_{(1)}, X_{(2)}, \ldots X_{(n)}$ olarak gösterirsek, gözlenen dağılım fonksiyonunu sembolik olarak aşağıdaki şekilde gösterebiliriz.

$$S_n(x) = \begin{cases} 0 & \text{eğer} \quad x < X_{(1)} \\ \dfrac{k}{n} & \text{eğer} \quad X_{(k)} \leqslant x \leqslant X_{(k+1)}\ ; \ k = 1, 2, \ldots n-1 \\ 1 & \text{eğer} \quad x \geqslant X_{(n)} \end{cases}$$

**1.2.1. K-S Uygunluk Testinin Varsayımları:**

Varsayımlarımızı aşağıdaki gibi sıralayabiliriz:

a) Alınan örnek rastgele seçilen bir örnektir.

b) Gözlenen örneklerin tam olarak belirlenmiş bazı sürekli dağılımlarından $[F_0(x)]$, gelip gelmediğini bu test ile anlamaya çalışırız. Bundan dolayı örnek dağılımının sürekli olduğunu varsayarız. Böyle durumlarda test tamdır. Fakat dağılımın belli parametreleri örnekten tahmin edildiği zaman, K-S uygunluk testi tam olarak uygulanamaz. En azından yaygın olarak tablolaştırılmış K-S test istatistik değerleri kullanılamaz. Eğer test bu durumlarda kullanılırsa tutucu bir teste dönüşecektir. Yani birinci tip hata işleme olasılığı ($\alpha$), K-S test istatistiğinin tablosu tarafından verilen değerden daha küçük çıkacaktır. Başka bir deyimle $H_0$ hipotezini reddetme olasılığı azalacaktır. Fakat $F_0(x)$'in dağılımı süreksiz ve $H_0$ hipotezi bu testle reddedilirse elde edilen sonuç çok güvenilirdir [18].

c) Hipotez edilen teorik dağılımın tam olarak örnekten bağımsız olarak belirlendiği varsayılır.

---

17 Jean Dickinson Gibbons, Nonparametric Statistical Inference, Mc Graw Hill, 1971, s. 73-74.

18 Malcolm, J. Slakter, "A Comparasion of the Pearson $X^2$ and Kolmogorov Goodness-of-fit Test with Respect to Validity", Journal of American Statistical Association, Vol. 60 (1965), s. 854-858.

d) Ölçme düzeyinin doğruluğunu sağlayabilmek için birbirine [19] eşit olan gözlem değerlerinin alınmamasına dikkat edilmelidir.

### 1.2.2. K-S Uygunluk Testinin Hipotezleri ve Test İstatistiği:

Tek ve çift taraflı hipotez testlerimizi aşağıdaki gibi belirtebiliriz:

Tek taraflı test:

$$A_1: \quad H_0 = F(x) \geqslant F_0(x) \quad ; \quad x \in (-\infty, \infty)$$
$$\quad\;\; H_1 = F(x) < F_0(x) \quad ; \quad \text{en az } x\text{'in bir değeri için}$$

veya

$$A_2: \quad H_0 = F(x) \leqslant F_0(x) \quad ; \quad x \in (-\infty, \infty)$$
$$\quad\;\; H_1 = F(x) > F_0(x) \quad ; \quad \text{en az } x\text{'in bir değeri için}$$

Çift taraflı test

$$H_0 : F(x) = F_0(x) \quad ; \quad x \in (-\infty, \infty)$$
$$H_1 : F(x) \neq F_0(x) \quad ; \quad \text{en az } x\text{'in bir değeri için}$$

Test istatistiklerimiz, $F_0(x)$ sürekli ve supremum $S_n(x)$ üzerinde bir atlama noktasına sahip değilse iki taraflı test için

$$D_n = \sup_{-\infty < x < \infty} |S_n(x) - F_0(x)|$$

tek taraflı ise, $A_1$'deki tek taraflı hipotezimiz için

$$D_n^+ = \sup_{-\infty < x < \infty} [F_0(x) - S_n(x)]$$

$A_2$'deki tek taraflı hipotezimiz için

$$D_n^- = \sup_{-\infty < x < \infty} [S_n(x) - F_0(x)]\text{' dir}^{20}.$$

$D_n$, $D_n^+$ ve $D_n^-$'in olasılık dağılımı, $F_0(x)$ sürekli olduğu sürece $F_0(x)$'e bağlı değildir. Bundan dolayı da $D_n$, $D_n^+$ ve $D_n^-$ test istatistikleri, bir serbest dağılım istatistikleri olarak bilinir [21].

Ayrıca $V_n = D_n^+ + D_n^-$ dönüşümü ile yeni bir test geliştirilmiş olup, bu test yardımı ile yayılmalarda ki noktaların yüzdesini test edebiliriz [22].

$S_n(x)$ daha önce de belirttiğimiz gibi şöyledir. Geçerli olan $X$'in herhangi bir değeri için

$$S_n(x) = \frac{k}{n}$$

---

19 Bradley, A.g.e., s. 300.

20 $F_0(x)$ süreksiz ve supremum $S_n(x)$ üzerinde bir atlama noktasına sahipse test istatistiklerimiz aşağıdaki gibidir:
$D_n = \max |S_n(x) - F_0(x)|$
$D_n^+ = \max [F_0(x) - S_n(x)]$
$D_n^- = \max [S_n(x) - F_0(x)] = -\max [F_0(x) - S_n(x)]$
Bradley, A.g.e. s. 296-298.

21 Gibbons, A.g.e., S. 76.

22 Urs V. Maag and Ghislaine Dicaire, "On Kolmogorov-Smirnov Type one Sample Statistics," Biometrica, Vol. 58 (1971), S. 653-656.

$k$: $x$'e eşit veya ondan daha küçük olan gözlemlerin sayısıdır. $x$'in her değerine göre $H_0$ hipotezi altında $S_n(x)$'in $F_0(x)$'e yakın olması beklenir. Eğer $H_0$ hipotezi doğru ise $F_0(x)$ ile $S_n(x)$ arasında hemen hemen hiçbir fark yoktur. Tersine bir durumda ise aralarında farklılık vardır.

### 1.2.3. K-S Uygunluk Testinde Karar Alma:

Hesapladığımız test istatistikleri $D_n$, $D_n^+$ ve $D_n^-$, $1-\alpha$ güven katsayısında K-S tablosundan bulunan değerlerden ($D_{n,\alpha}$, $D^+_{n,\alpha}$ ve $D^-_{n,\alpha}$) büyük çıkarsa, $H_0$ hipotezimizi $\alpha$-anlamlılık seviyesinden reddederiz.

Çift taraflı test hipotezimiz için örnek sayısı yirmiye eşit veya daha küçük olursa tamdır.

Tablo değerleri çok iyi bir formüle dayandırıldığı için örnek sayısının 20 den büyük olduğu durumların bazılarında da tamdır.

K-S uygunluk testinin tablo değerleri, $1 \leqslant n \leqslant 40$ için verilmiştir. Eğer $n > 40$ ise bunun için $D_{n,\alpha}/\sqrt{n}$ yaklaşımı kullanılabilir [23].

### 1.2.4. K-S Uygunluk Testinin Etkinliği:

Herşeyden önce aklımıza şöyle bir soru gelebilir, K-S tek örneklem uygunluk testi süreksiz verilere de aynı etkinlikte uygulanabilir mi? Yukarıda da değindiğimiz gibi K-S uygunluk testi sürekli dağılım fonksiyonlarını varsayarak yola çıkar. Eğer test süreksiz dağılım fonksiyonlarını kullanırsa tutucu bir teste dönüşür.

Ramachandramurty, tek ve iki örneklemli durumlarda K-S testinin etkinliğinin alt limit değerini hesaplamıştır. Bu hesaplamalara dayanarak K-S testinin $t$-testine göre A R E'si için bir alt limit tarifi vermiştir [24].

Conover de, hipotez edilen dağılım fonksiyonu süreksiz olduğu zaman, K-S tek örneklem uygunluk test istatistiğinin tam olarak kritik düzeyinin bulunması için bir yöntem geliştirmiştir. Örnek sayısı 30'a eşit veya daha küçük olursa, bu yöntem tam olarak aynı doğruluk oranında araştırmaya uygulanabilmektedir [25].

Massey ise örnek sayılarımızın çok büyük olduğu durumlarda K-S testinin gücünün,

$$1-(2\pi)^{1/2}\int_{2[\sqrt{n}-D_{n,\alpha}(n)]}^{2[\sqrt{n}+D_{n,\alpha}(n)]} e^{-z^2/2}\,dz$$

integrali ile bulunan değerden daha küçük olmayacağını göstermiştir. Fakat buna rağmen yukardaki formül yardımı ile bulunan değerin, gerçek K-S testinden daha küçük olacağı da belirtilmiştir [26].

Lee, S.W, normal dağılım testi ile K-S testinin gücünü karşılaştırmış ve bazı durumlarda bu uygunluk testinin daha güçlü olduğunu belirtmiştir. Uygun olma-

---

23 Wayne W. Daniel: Applied Nonparametric Statistics, Haughton Mifflin Co., 1978, S. 462.

24 Ramachandramurty, P.V.: "On the Pitman Efficiency of onesided Kolmogorov-Smirnov Test for Normal Alternatives", Annals of Mathematical Statistics, Vol. 37 (1966), S. 940-944.

25 William Jay Conover: "A Kolmogorov Goodness-of-fit Test for Distributions" Journal of American Statistical Association, Vol 67 (1972), s. 591-596.

26 Frank J. Massey: "The Kolmogorov-Smirnov Test for Goodness-of-fit" Journal of American Statistical Association, Vol 46 (1951), s. 68-78.

yan durumlarda bile K-S uygunluk testi normal testden çok daha kötü çıkmamıştır [27].

K-S testinin parametrik testlerle karşılaştırmasını-değişik $\alpha$ ve örnek sayılarında yapmak için oldukça yoğun çalışmalar günümüzde de üzerinde durulan bir konudur.

**2- İki Uygunluk Testinin Karşılaştırılması:**

K-S ve Ki-Kare uygunluk tek örneklem testleri birbirlerinin seçeneğidirler. Bunların uygulanabildiği veri ve türlerine başka bir testle yaklaşma olanağı yoktur. İki uygunluk testinin karşılaştırılmasına ve değişik durumlarda hangisinin seçilebileceğine geçmeden önce, bu testlerin birinden diğerine nasıl ve hangi durumlarda geçildiğine kısaca değinelim.

Eğer $F_0(x)$ herhangi bir sürekli dağılım fonksiyonu ise, sıfıra eşit veya büyük bütün $Z$ değerleri ($Z \geqslant 0$) için, aşağıdaki limit hali gerçektir.

$$\underset{n \to \infty}{\text{Limit}} \quad P\left(D_n^+ < \frac{Z}{\sqrt{n}}\right) = 1 - e^{-2z^2}$$

Bu limitin sonucu olarak, eğer $F_0(x)$ herhangi bir sürekli dağılım fonksiyonu ise sıfıra eşit veya büyük bütün $Z$ değerleri ($Z \geqslant 0$) için, $n$ sonsuza giderken $4n\,(D_n^+)^2$'-nin limit dağılımı iki serbestlik derecesine sahip bir Ki-Kare dağılımıdır [28]. Başka bir ifade ile

$$\underset{n \to \infty}{\text{Limit}} \quad P\,[4n(D_n^+) < k] = X^2_{(2)} \; ; \text{ Bütün k'lar} \geqslant 0$$

K-S uygunluk testi, örneğin rastgele ve birbiri içinde yakın ilişkiler olduğunu kabul edip, örnek yığının sürekli olduğunu varsaymasına rağmen Ki-Kare Uygunluk testi için bunlar önemli değildir. Ayrıca Ki-Kare uygunluk testi örnek yığınını süreksiz ve sınıflayıcı verilerle kullanıldığı zaman uygundur ve böyle verilere uygulamada sık sık rastlanır.

Kuramsal olarak K-S uygunluk testi sonsuz gözlem sayısı ihtiva ettiği halde, Ki-Kare sınırsız sınıf sayısı içerir. Fakat uygulamada her iki durum da söz konusu değildir. Genelde Ki-Kare uygunluk testi yalnızca çok büyük ve her sınıftaki sıklıkların beklenen değerlerinin çok küçük olmadığı durumlarda, K-S uygunluk testi ise her türlü örnek büyüklüğü (özellikle küçük olanlar tercih edilir) için uygulanabilir [29].

K-S uygunluk testi incelenen değişkenin sürekli bir dağılım olduğunu varsayar. Hipotez edilen birikimli dağılım $F_0(x)$ süreksiz olduğunda, tahmin ettiğimiz örnek dağılımı ($\hat{D}_n$), K-S test istatistiğimiz $D_n$ den farklı olacaktır. Çünkü $D_n$ kritik değerleri sürekli durumlar için tablolaştırılmıştır. Bu durumda test kullanılırsa sonuç olarak çıkan olasılık ifadesindeki hata hemen hemen kesindir (Test tutucudur). Lilliefors, hipotez edilen dağılım normal veya exponensiyal olduğu zaman, $\hat{D}_n$'in kritik değerleri için tam olarak tablolar oluşturmuştur. Bu durumlarda K-S uygunluk testinin, Ki-Kare testinden daha güçlü çıktığını örneklerle göstermiş ve K-S uygunluk

27 Conover, A.g.e., S. 300-301.
28 Gibbons, A.g.e. S. 83.
29 Aynı eser, S. 75.

testinin seçilmesini belirtmiştir [30].

K-S uygunluk testi iki taraflı testler için kullanıldığı gibi bu testle, sapmaların yönüde $D_n^+$ ve $D_n^-$ test istatistiklerinden kolaylıkla bulunabilir. Ki-Kare uygunluk testi gözlenen ile beklenen değerler arasındaki uyuşmanın yönünü gösteremez. Çünkü Ki-Kare testindeki sapmalar kareli bir yapıya sahiptir. K-S test istatistiği aynı zamanda gerçek yığın dağılımının güven sınırları tahmininde ve belirlenmiş bir varsayım için istenen minimum örnek büyüklüğünün tahmin edilmesinde de kullanılabilir. Ki-Kare test istatistiğinin bu amaçlar için kullanılabilecek benzer özellikleri yoktur [31].

Ki-Kare test istatistiği faydalı gelebilen bir özelllik için anlamlıca toplanabilir, halbuki K-S uygunluk testi her gözlemi ayrı ayrı ele alıp inceler. Böylece K-S uygunluk testinin, Ki-Kare testinde uygulanan, sınıflamaların birleştirilmesinden doğan bilgi kaybı yoktur ve bu durumlarda daha güçlüdür.

Ki-Kare uygunluk testi daha önce değindiğimiz gibi sınıflamaların sayısı azaltılarak serbestlik derecesi küçültülür ve parametrelerin örnekten tahmin edilmesini gerektiren durumlarda kolaylıkla kullanabilecek bir test durumuna sokulur. K-S testinin bu türden bilinen kolaylıkları yoktur. Şüphesiz Ki-Kare uygunluk testinin sınıflamaların birleştirilmesinden doğan bir sakıncası da vardır. Ki-Kare uygunluk testinde sınıflamalar birleştirildiği zaman, Ki-Kare test istatistiğinin hesaplanmış değeri — bütün örnek büyüklükleri için — tek değildir. Bundan dolayıda bu test sınıflamaların etkisi altında kalır ve tahmini bir testtir. Buna karşılık K-S uygunluk testi küçük örnek büyüklükleri için bile tam olup, büyük örnek durumlarında da tahmin derecesi çok kolaydır.

K-S uygunluk testinde elde edilen verilerin değerlendirilmesini yapmak Ki-Kare uygunluk testinden daha kolaydır. Ama bu üstünlüğü günümüzdeki teknolojik gelişmenin çok hızlı olmasından dolayı kaybolmuştur.

Bu iki testin güç etkinliğinin karşılaştırılması ve hangisinin seçileceği konusunda bugüne kadar belli bir temel oluşturulamamıştır. Böyle bir temel olmamasının zorluğu bu iki uygunluk testinin güç fonksiyonlarının değişik niceliklere bağlı olmasından dolayıdır [32].

---

30 Hubert W. Lilliefors: "On the Kolmogorov-Smirnov Test Normality Distribution With Mean and Variance Unknown" Journal of American Statistical Association, Vol. 64 (1969), S. 399-402.

................., "On the Kolmogorov-Smirnov Test for the Exponential Distribution with Mean Unknown," Journal of American Statistical Association, Vol. 64 (1969), S. 387-389.

31 K-S testinin güven aralığının oluşturulması için bakınız: Bradley, A.g.e., S. 301-302.

32 Gibbons, A.g.e., S. 87.
K-S uygunluk testinin gücü,
Sup $|F_0(x) - F(x)|$'e bağlı olduğu halde
$-\infty < x < \infty$
Ki-Kare uygunluk testinin gücü

$$\sum_{i=1}^{k} \frac{[F_0(a_i+1) - F_0(a_i)] - [F(a_i+1) - F(a_i)]}{F(a_i+1) - F(a_i)}$$

kavramına bağlıdır. Formülde
$k$: sınıflama sayısı
$a_i$: sınıfların sahip olduğu limit değerlerdir.

Bu iki uygunluk testinden hangisinin seçilebileceği testin tam olup olmadığına, kullanılabilir olup olmadığına ve büyük ölçüde testi uygulayacak olan araştırmacıya kalmıştır. K-S uygunluk testi gerekli olan bütün parametreleriyle sürekli bir yığını içeriyorsa tamdır. Buna karşılık Ki-Kare uygunluk testinin her zaman yaklaşık bir değer olduğu hatırlanmalıdır. K-S uygunluk testi örnek büyüklüğüne bağlı kalmaksızın (küçük örnek büyüklükleri özellikle tercih nedenidir) ve varsayımları sağlandığı durumlarda güvenle kullanılabilir.

## FAYDALANILAN KAYNAKLAR


Bradley, James V., Distribution-Free Statistical Test, Prentice Hall Inc., New Jersey, 1968.

Conover, William Jay, Practical Nonparametric Statistics, John Wiley, New York, 1971.

.................., "A Kolmogorov Goodness of-fit Test for Discontinous Distribution", Journal of American Statistical Associotion, Vol 67 (1972), S. 591-596.

Daniel, Wayne W., Applied Nonparametric Statistics, Haughtan Mifflin Co., 1978.

Freund, F.E. and Williams, F.J., Elementary Business Statistics, Prentice-Hall, 1969.

Gibbons, Jean Dickinson, Nonparametric Statistical Inference, Mc Graw Hill, 1971.

Harris, Bernard, Theory of Probability, Addison-Wesley, 1966.

Hoggy, Robert, V. and Craig, Allen T., Introduction to Mathematical Statistics, Mc Millan, Hong-Kong, Third Edition, 1972.

Lehmann, E.L., Nonparametrics, Holden Day, New York, 1975.

Lilliefors, W. Hubert, "On the Kolmogorov-Smirnov Test Normality Distribution with Mean and Variance Unknown", Journal of American Statistical Association, Vol 64 (1969), S. 399-402.

.................., "On the Kolmogorov-Smirnov Test for the Exponential Distribution with Mean Unknown" Journal of American Statistical Association, Vol. 64 (1969), S. 387-389.

Maag, Urs, V. and Dicaire, Ghislaine, "On Kolmogorov-Smirnov Type One Sample Statistics", Biometrica, Vol 58 (1971), S. 653-656.

Massey, Frank, J., "The Kolmogorov-Smirnov Test for Goodness-of-Fit", Journal of American Statistical Association Vol. 46 (1951), S. 68-78.

Ramachandramurty, P.V., "On the Pitman Efficiency of one sided Kolmogorov-Smirnov Test for Normal Alternatives", Annals of Mathematical Statistics, Vol. 37 (1966), S. 940-944.

Slakter, Malcolm, J., "A Comparison of the Pearson $X^2$ and Kolmogorov Goodness of-Fit. Test for Small but Equal Expected Frequincies with Respect to Validity", Journal of American Statistical Association, Vol. 60 (1965), S. 854-858.

Yamane, Taro, Statistics and Introductory Analysis, Harper Enternational Edition, Singapore, 1973.

Yoğurtçugil, M. Kemal, Ki-Kare Üzerine Bir Deneme, İstanbul Üniversitesi İktisat Fakültesi, İstanbul, 1978.