SERİ SERIE B | CİLT TOME XXI | SAYI FASCICULE 1 | 1971

İSTANBUL ÜNİVERSİTESİ

ORMAN FAKÜLTESİ DERGİSİ

REVUE DE LA FACULTÉ DES SCIENCES FORESTIÈRES
DE L'UNIVERSITÉ D'ISTANBUL

İSTATİSTİKSEL TESTLER: HANGİSİ, NE ZAMAN?

Yazan :

Dr. H. Alptekin GÜNEL

G İ R İ Ş

İstatistiğin ilmi araştırmalardaki rolünü üç grupta toplamak mümkündür: Donelerin toplanması ve tanımlama, analiz ve takdir.

Tanımlamadan maksat, anlam ve kapsamını kaybetmeden, toplanan donelerin miktarını asgariye indirmek ve toplumu karakterize eden değerleri, yani parametreleri ortaya koymaktır. İstatistiğin bu fonksiyonu belirli bilim dal veya dallarına münhasır değildir. «Ortalama değer» kavramı günlük hayatta dahi en çok kullanılan terimlerden biridir. Ekstrem değerlerin uyarabileceği yanlış kanıları önlemek için «ortalama olarak» ifadesi sık sık kullanılmaktadır. Bu şekli ile, «genellikle» deyimi «ortalama» ile aynı anlamdadır. Bu kullanış şekli hatalı olarak nitelendirilemez. Zira, diğer özelliklerine ilâveten, bir toplumu en iyi özetleyen kavram, ortalama kavramıdır. Bununla beraber, ortalama, bir toplumu bütün özellikleri ile ortaya koymağa yeterli değildir. Ortalamaya ek olarak, diğer bir ifadeye ihtiyaç vardır. «Değişkenlik» kavramı bu amaçla ortaya atılmıştır. Ancak, değişkenlikten ne anlaşıldığı her zaman açık değildir. Bunlar arasında, ortalamadan farkların kareleri ortalaması, daha alışılmış ifadesi ile varyans, bu amaçla en çok kullanılan, bu nedenle de standart bir ifade olarak itibar edilen bir kavram olmuştur.

İstatistikte diğer önemli bir kavram «korelasyon»dur. İki veya daha fazla değişken arasındaki bağlılık derecesini ifade etmektedir. Bu bağımlılık, matematiğin fonksiyonel bağımlılığından, mahiyet itibariyle farklıdır. Hatta, çok kere böyle gerçek bir ilişki olmayabilir. Meselâ, tabii olarak kabul edilen ve bir çok çalışma ve metodlara baz olarak alınan, tek ağaçta çap - boy ilişkisi bu türden bir bağımlılıktır. Her iki değişkenin de zamanın bir fonksiyonu olması, bu iki eleman arasında da bir korelasyonun gözükmesine yol açmaktadır. Esas itibariyle, bu korelasyon insan kontrolu dışındaki etkenlerin hakimiyeti altındadır ve bu nedenle de ancak istatistiki anlama bir ilişkiden söz edilebilmektedir.

Yukarda sözü edilen «istatistiklerin» ortaya konabilmesi, elde ölçme ve gözlem sonuçlarının bulunmasını gerektirir. Ancak, bu sonuçlardan geçerli yargılara varılabilmesi, onların belirli yapıda ve belirli esaslara göre derlenilmiş olmasını gerektirir. Burada önemli olan husus, bu esasların gene istatistik bilimi tarafından, bütün ayrıntıları ile ortaya konabilmiş olmasıdır. Söz konusu esasların ne olduğu, araştırmanın amacına, eldeki mali ve teknik imkânlara bağlıdır. Bu esaslar, genellikle bir ihtiyacın sonucu olarak geliştirildiklerinden, diğer bir etken olan «uygulanabilirlik» çok kere ya hiç önemli olmamakta veya tali derecede kalmaktadır.

Bilimsel araştırmaların amacı, belirli bir varsayımın kontroludur. İstatistik biliminin ikinci fonksiyonu bu durumda ortaya çıkmaktadır.

Analize tabi tutulacak varsayımın açık ve kesin olması gerekir. İstatistiki anlamda yapılacak varsayım kontrolunda veya testinde, varsayımın ancak yanlışlığı gösterilebilir. Zira, bu türden kontrollerde, sonlu sayıda yapılacak denemelerle bir varsayımın doğruluğu ispatlanamaz.

Yukarda ifade edilen ve çeşitli şekilde hesaplanan istatistikler, böyle bir kontrolun ancak birer «malzemesi»dirler. Söz konusu kontrollerde kullanılmayacak istatistikler «gösteriş için girişilmiş, meyvasız gayretlerden» ileri gidemezler.

İstatistiksel yargıların ana metodunun tümevarım olduğu göz önünde tutulursa, böyle bir yargıda daima bir belirsizlik payı bulunacaktır. Bu sonuç metodun karakteri icabı olduğu kadar, istatistik bilimine konu olan olayların niteliklerinin de zorunlu kıldığı bir vakıadır. Diğer taraftan, bu belirsizlik tamamen yok edilemezse bile, istenen sınırlar içinde tutulabilir.

Bu yazıda, varsayım kontrolerinin (testlerinin) teorik tartışılmasından ziyade, istatistiksel kontrollerin uygulanmaları ile ilgili hususlar üzerinde durulacak ve bu amaçla en çok kullanılan parametrik istatiksel kontrollerin takdimine çalışılacaktır. Buna paralel olarak da, yazının kapsamına uygun temel kavramların formel olmayan tanımları verilecektir. Böylece, mahiyet itibariyle aynı, fakat kullanılmaları daha karışık görünen metodların tartışılmalarına zemin hazırlanmış olacağı düşünülmüştür.

İstatiksel Testlerle İlgili Bazı Kavramlar

Yukarda, istatistiksel yargıların, genellikle tümevarım metodunu kullandıkları, bu nedenle varılan yargılarda, daima bir belirsizlik payının bulunduğu ifade edilmişti. Söz konusu belirsizlik, ölçme ve gözlemlerde tesadüfiyeti emniyet altına almak, ölçme ve gözlem sayısını arttırmak ve daha çok sayıda deneme yapmakla, tamamen ortadan kaldırılmasa bile, istenen sınırlar içinde tutulabilir. Bu şekilde tayin edilen belirsizlik miktarına «Güvenirlik Derecesi» (Level of Significance) denir. Araştırmanın amacına bağlı olmakla beraber, en çok kullanılan güvenirlik dereceleri % 10, % 5 veya % 1 dir. Burada önemle belirtilmek istenen husus, güvenirlik derecesini seçerken, günlük hayattaki güvenirlik ile istatistiksel güvenirliğin aynı olmadığını, problemin karekterine göre güvenirlik derecesinin farklı alınabileceğini göz önünde tutmak gerektiğidir.

İstatistiksel yargıların yapısındaki belirsizlik nedeni ile, gerçekte doğru bir varsayımı red etmek veya yanlış bir varsayımı doğru olarak kabul etmek mümkündür. Birinci halde yapılan hataya I. tip hata; ikinci halde yapılan hataya ise II. tip hata denir (Tablo 1).

Tablo 1. Varsayım Kontrolundaki hata çeşitleri

	Kabul	*Red*
Doğru Varsayım	Doğru yargı	I. Tip hata
Yanlış Varsayım	II. Tip hata	Doğru yargı

I. tip hata ihtimalini (α), II. tip hata ihtimalini (β) ile göstermek mutat hale gelmiştir. I. tip hatanın miktarı araştırıcı tarafından tesbit edilebilmesine karşılık, bu imkân II tip hata için sınırlıdır. Mamafih, II. tip hata I. tip hata ile sıkı sıkıya bağlıdır. Birisinin miktarında yapılacak bir azaltma diğerinin miktarında bir yükselme meydana getirir. Hangi tip hatanın daha küçük tutulacağı konusu, araştırmanın mahiyetine bağlıdır. Meselâ tehlikeli bir ilâcın tehlikesiz olarak kabulü (II. tip hata), zararsız bir ilâcın zararlı olarak kabul edilmesinden daha önemlidir. Bu taktirde ikinci tip hatayı küçük tutmak, buna mukabil daha büyük bir I. tip hata miktarı ile yetinmek gerekir. Yanlış bir varsayımı red etme ihtimaline ($1 - \beta$) istatiksel testin kuvveti denir. Bu değeri her zaman hesaplamak mümkün olmamaktadır.

Bir istatiksel testlerde, genellikle üç durum söz konusudur :

1 — İstatiksel teste konu olan toplum parametresi belirli bir değere eşittir,

2 — İlgili parametre belirli bir değere eşit veya ondan büyüktür,

3 — İlgili parametre belirli bir değere eşit veya ondan küçüktür.

Bu üç varsayımın her biri, kendi hali için, sıfır varsayımı olarak adlandırılır ve H_0 ile gösterilir. Bu üç sıfır varsayımın karşıt varsayımları H_A (alternatif varsayımlar) sırasiyle şöyledir :

1.1 Toplumun parametre değeri belirli değere eşit değildir,
2.1 Toplumun parametre değeri belirli değerden küçüktür,
3.1 Toplumun parametre değeri değerden büyüktür.

Toplumun parametre değeri, belirtilen değere eşit değilse, ondan ya büyük veya küçüktür. Bu nedenle, böyle bir teste iki taraflı (two sided) test denir. Diğer hallerde, parametre değeri, belirtilen değerin ancak bir tarafından bulunabilir. Bu taktirde, tek taraflı istatistiksel test yapılacak demektir.

İstatistiksel testler

İstatistiksel testleri iki genel grupta toplamak mümkündür :

A — Ortalamalarla ilgili testler,

B — Varyanslarla ilgili testler.

A — Ortalamalarla ilgili testler

Bu tür testlerde şu durumlar söz konusudur :

Aa. Tek bir ortalama ile ilgili testler,

Ab. İki ortalama ile ilgili testler,

Ac. Eşlendirilmiş gözlemlerle (paired observations) ilgili testler,

Ad. İkiden fazla ortalamalarla ilgili testler.

B — Varyanslarla ilgili testler

Ba. Tek bir varyansla ilgili testler,

Bb. İki varyansla ilgili testler,

Bc. İkiden fazla varyansla ilgili testler.

Yukardaki her bir hal için kullanılacak istatistikler ve bu amaçla yapılması gereken işlemler aşağıda özetlenmeye çalışılmıştır.

A — Ortalamalarla İlgili Testler

Aa. Tek bir ortalama ile ilgili testler

Toplum ortalamasının önceden verilen belirli bir değere eşit olduğu varsayımının testidir. Ortalamanın ait olduğu toplumun varyansının önceden bilinmesi halinde, kullanılacak istatistik Z — istatistiğidir.

$$Z = \frac{\overline{X} - \mu_0}{\sigma / \sqrt{N}} \quad (1)$$

Eşitlikte,

$\overline{X}$ = Örnekten hesaplanan ortalama,

μ_0 = Toplum ortalamasının eşit olduğu varsayılan değer,

σ^2 = Toplumun bilinen varyansı

σ = Varyansın kare kökü veya standart sapma,

N = Örnekteki ünite sayısı'nı göstermektedir.

Varsayımın testinde takip edilen hareket tarzı şöyledir :

1 — Sıfır hipotezi ifade edilir,

$$\mu = \mu_0$$

2 — Güvenirlik derecesi (α) yani I. tip hata yapma ihtimali, seçilir,

3 — (1) eşitliğinden Z — istatistiği hesaplanır. Hesaplamaya esas olan örnek, normal dağılımlı bir toplumdan tesadüfi olarak alınmışsa, Z — değeri standart normal dağılımlıdır.

4 — Seçilen güvenirlik derecesine göre, normal dağılım tablosundan kritik Z — değerleri bulunur. Kritik değerler, bu değerlerden büyük veya küçük olma ihtimaleri, seçilen güvenirlik derecesi kadar olan değerlerdir. Bu değerlerin sınırladıkları bölgeye «kritik bölge» denir. Meselâ, normal dağılım ve $\alpha = \%\,5$ için, kritik Z — değerleri, iki taraflı test halinde, $\pm$ 1,96 dır.

5 — (1) eşitliği ile bulunan değer, kritik değerlerden büyük veya küçükse, yani kritik bölgeye düşüyorsa, $U = U_0$ varsayımı red edilir. Aksi halde, varsayımı red edecek yeterli delil bulunamamıştır.

Sıfır varsayımının, toplumun aritmetik ortalamasının belirli bir μ_0 değerinden büyük veya eşit hali için hareket tarzı yukardakinin tamamen aynıdır. Ancak, test tek taraflı olduğundan, kritik değer $+ 1,65$ dir.

Sıfır varsayımının, toplumun aritmetik ortalamasının belirli bir μ_0 değerine eşit veya ondan küçük hali için de test tek taraflı olup kritik değer $-1,65$ dir.

Toplumun varyansı bilinmiyorsa, bu taktirde, ortalama ilgili varsayımın testinde kullanılacak istatistik

$$t = \frac{\overline{X} - \mu_0}{s / \sqrt{N}} \tag{2}$$

eşitliğinden hesaplanır. (2) nolu eşitlikteki terimler, eşitlik (1)'dekilerle aynı olup, sadece (S) örnekten hesaplanan standard sapmadır. Varsayımın testi yukardaki ile aynıdır. Ancak, eşitlik (2) t— dağılımı gösterdiğinden, kritik değerlerin bulunması biraz değişiktir. Zira, t— değeri, seçilen güvenirlik derecesinin bir fonksiyonu olduğu kadar serbestiyet derecesine de bağlıdır. Bu nedenle, kritik değerlerin tayininde, güvenirlik derecesinin yanında, serbestiyet derecesinin de bilinmesine ihtiyaç vardır. Yukardaki eşitlikte, serbestiyet derecesi ölçme sayısının bir noksanı, yani $(N - 1)$'dir. Meselâ, $N = 25$, $\alpha = \%\,5$ ve iki taraflı test için, serbestiyet derecesi $(25 - 1 = 24)$ olup, kritik t— değerleri $\pm 2,064$; buna karşılık, aynı serbestiyet derecesi ve α değeri, fakat tek taraflı test için t— kritik değeri 1,711 dir.

Ab. İki ortalama ile ilgili varsayımlar

İki toplum ve bu toplumların U_1 ve U_2 gibi ortalamaları ile σ_1^2 ve σ_2^2 gibi varyansları olduğunu ve bu varyansların bilindiğini kabul edelim. Genellikle, teste tabi tutulmak istenen varsayım, bu iki toplumun ortalamalarının eşit olup olmadığıdır. Söz konusu toplumların ortalamaları eşitse, $\overline{X}_1$ birinci toplumdan alınan örnekten bulunan ortalamayı, $\overline{X}_2$ ikinci toplumdan alınan örnekten bulunan ortalamayı göstermek üzere $U_1 - U_2 = 0$ olacağından,

$$Z = \frac{(\overline{X}_1 - \overline{X}_2) - (U_1 - U_2)}{\sqrt{\dfrac{\sigma_1^2}{N_1} + \dfrac{\sigma_2^2}{N_2}}} = \frac{(\overline{X}_1 - \overline{X}_2)}{\sqrt{\dfrac{\sigma_1^2}{N_1} + \dfrac{\sigma_2^2}{N_2}}} \tag{3}$$

eşitliği, toplumlar normal dağılımlı veya N'ler yeter derecede büyükse, standart normal dağılım gösterecektir.

Daha özel bir hal olarak, toplumların varyansları eşit olabilir. Bu takdirde, (3) eşitliği

$$Z = \frac{(\bar{X}_1 - \bar{X}_2)}{\sigma \sqrt{\frac{1}{N_1} + \frac{1}{N_2}}} \tag{3a}$$

şeklinde yazılabilir.

İki toplumun varyansları biliniyor ve bu varyanslar eşitse, ortalamaların eşitliği ile ilgili varsayımın testinde aşağıdaki şekilde hareket edilir :

1 — Varsayım ifade edilir,

$$U_1 = U_2$$

2 — Güvenirlik derecesi seçilir,

3 — (3a) eşitliğinden Z — değeri hesaplanır. Yapılacak varsayım testi iki taraflıdır. Bu nedenle, seçilen güvenirlik derecesine göre, kritik değerler dolayısiyle, kritik bölge tespit edilir,

4 — Hesaplanan Z — değeri kritik bölge içine düşüyorsa varsayım red edilir. Aksi halde, varsayımı red edecek yeterli delil mevcut değildir.

Ortalamalar ile ilgili varsayım, ortalamalardan birinin diğerinden küçük olduğu şeklinde olabilir. Bu durumda, kullanılacak istatistik gene (3a) eşitliği ile bulunacak değerdir. Ancak, test tek taraflıdır, dolayısiyle tek bir kritik bölge vardır.

Toplumların varyansları biliniyor, fakat bu varyanslar eşit değilse, varsayımın testinde yukardaki gibi hareket edilir. Kullanılacak istatistik eşitlik (3) yardımı ile bulunur.

Toplumun varyanslarının bilinmemesi halinde, kullanılacak istatistik

$$\frac{\bar{X}_1 - \bar{X}_2}{S_p \sqrt{\frac{1}{N_1} + \frac{1}{N_2}}} \tag{4}$$

eşitliği ile bulunur. Eşitlikteki S_p^2 «birleştirilmiş varyans» olup

$$S_p^2 = \frac{(N_1 - 1)\, S_1^2 + (N_2 - 1)\, S_2^2}{N_1 + N_2 - 2} \tag{5}$$

şeklinde hesaplanır. Eşitlik (5)'deki

S_1^2 = Birinci toplumdan alınan örnekten hesaplanan varyans,

S_2^2 = İkinci toplumdan alınan örnekten hesaplanan varyans'tır.

(4) eşitliği t — dağılımı gösterir. Dağılımın serbestiyet derecesi $N_1 + N_2 - 2$ dir. Yukardaki varsayımın testi iki taraflıdır.

Ortalamalardan birinin diğerinden küçük (veya büyük) olduğu varsayımı için, gene (4) eşitliği ile hesaplanan t — istatistiği kullanılır. Bu kez test tek taraflıdır.

Ortalamalarla ilgili varsayımın testinde, toplumların varyansları bilinmiyorsa, bu taktirde kullanılacak istatistik

$$\frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\dfrac{S_1^2}{N_1} + \dfrac{S_2^2}{N_2}}} \tag{6}$$

dır. (6) eşitliği yaklaşık olarak t — dağılımı göstermekte olup, dağılımın serbestiyet derecesi

$$f = \frac{\left(\dfrac{S_1^2}{N_1} + \dfrac{S_2^2}{N_2}\right)}{\dfrac{\left(\dfrac{S_1^2}{N_1}\right)^2}{N_1 - 1} + \dfrac{\left(\dfrac{S_2^2}{N_2}\right)^2}{N_2 - 1}} - 2$$

yardımı ile bulunur. f — değeri tam sayı olarak bulunamazsa, t — tablosunda enterpolasyon yapılmalıdır. Örnekler yeter derecede büyükse, en yakın tam sayıya yuvarlama yapmak önemli bir hata doğurmaz.

Ac. Eşlendirilmiş gözlemlerle ilgili testler

İki toplumun varyanslarının eşit olmadığı kanısı mevcut ve her iki toplumdan alınan örnekler arasında bir korelasyon varlığından şüphe ediliyorsa, bu taktirde, iki ortalamanın eşitliği testinin, eşlendirilmiş gözlemler yardımı ile yapılması daha uygundur.

X_{i1} 1'inci toplumdan alınan i'inci ölçme i = 1,2, ..., N

X_{i2} 2'nci toplumdan alınan i'inci ölçme

$d_i = X_{i1} - X_{i2}$

$\bar{d} = d_i$ lerin ortalaması

S_d^2 = d_i lerin varyansı'nı göstersinler.

Bu takdirde, takip edilecek hareket tarzı şöyledir :

1 — Varsayım ifade edilir,

$U_1 = U_2$

2 — Güvenirlik derecesi seçilir,

3 — $t = \dfrac{\bar{d}}{S_d / \sqrt{N}}$ istatistiği hesaplanır. (9)

Varsayımın testi tek taraflıdır.

Ad. İkiden fazla ortalamalarla ilgili varsayım testi

İkiden fazla ortalamalarla ilgili varsayımlarının testi varyans analizinin konusudur. Varyans analizinin çok değişik hallere uygulanabilmesi nedeni ile, prensip itibariyle aynı olmakla beraber, hesap ayrıntıları farklı şekilleri vardır. Söz konusu bütün bu hallerin burada tartışılması yazının maksat ve çerçevesinin dışında kalmaktadır. Bu nedenle, burada en basit hal olan tek girişli (one - way) varyans analizi modeli kısaca verilmeye çalışılacaktır.

Yukarda açıklanan sınırlamalar karşısında, üzerinde durulacak iki varsayım ve bunların analizi şöyledir :

— Bütün ortalamalar birbirine eşittir

Bu varsayımın testi aşağıdaki şekilde yapılır :

1 — Varsayım ifade edilir,

$U_1 = U_2 = \ldots\ldots = U_p$

2 — Güvenirlik derecesi seçilir,

3 — $F = \dfrac{\text{Gruplar arası varyans}}{\text{Gruplar içi varyans}}$ (10)

istatistiği hesaplanır.

4 — Örnekler, normal dağılımlı toplumlardan alınmış veya yeter derecede büyükse; ayrıca, örneklerin alındığı toplumların varyansları eşitse, (10) eşitliği (p — 1) ve (ΣN_i — p) serbestiyet dereceli F — dağılımı gösterir. Test tek taraflıdır.

5 — Seçilen güvenirlik derecesi ve serbestiyet derecelerine göre, F — dağılım tablosundan kritik değer tespit edilir. Eşitlik (10) ile hesaplanan değer tablo değerinden büyükse varsayım red edilir.

— Ortalamalar Arasında Doğrusal Bir İlişki Vardır *

Herhangi bir ortalama (veya ortalamalar) diğerlerinin doğrusal bir fonksiyonu olarak varsayılabilir. Meselâ, birinci toplumun ortalaması diğer toplumların ortalamalarının toplamının üçte ikisine eşittir şeklinde bir varsayım söz konusu olabilir. Buna göre

$$U_1 = \frac{2}{3} (U_2 + U_3 + \ldots + U_p) \quad \text{veya}$$

$$3\, U_1 — 2\, (U_2 + U_3 + \ldots + U_p) = 0$$

yazılabilir.

Daha genel olarak yukardaki varsayım

$$a_1 U + a_2 U_2 + \ldots\ldots\ldots + a_p U_p = 0$$

şeklinde gösterilebilir. Bu taktirde

$$F = \frac{\left(\sum_{i=1}^{p} a_i \bar{X}_i \right)^2}{S^2 \sum_{i=1}^{p} a_i^2 / N} = \frac{\left(\sum_{i=1}^{p} a_i T_i \right)^2}{NS^2 \sum_{i=1}^{p} a_i^2} \qquad (11)$$

eşitliğinden yararlanılır. (11) eşitliği F — dağılımı gösterir. Eşitlikteki

(*) Birinci varsayım, aslında, bu varsayımın özel bir halidir : $U_1 = U_2 = \ldots\ldots = U_p$ eşitliğinden $U_1 = U_2$; $U_1 = U_2$; ....; $U_1 = U_p$ yazılabilir. Burada ise, $(P^{-1})\, U_2 — U_2 \ldots\ldots — U_p = 0$ elde edilir. Bu, ortalamalar arasında doğrusal bir ilişkiyi gösterir. Ancak, bu durumda, ortalamaların katsayıları daima (P^{-1}), — 1, ——, — 1 dir. Bu nedenle iki varsayımı ayrı mütalâa etmek daha uygun bulunmuştur.

a_i = i inci ortalamaya ait katsayı,

T_i = i inci örnek toplamı.

N = Örnekteki ünite sayısı, (bütün örneklerde aynı alınmıştır.)

S^2 = Gruplar içi varyans'ı

göstermektedir. (11) eşitliği ile bulunan F — istatistiğinin serbestiyet dereceleri 1, $(\Sigma N_i - P)$ dir.

$$Q^2 = \frac{\left(\sum_{i=1}^{p} a_i T_i\right)^2}{N \sum_{i=1}^{p} a_i^2} \tag{12}$$

ifadesine münferit serbestiyet derecesi (individual degrees of freedom) denir.

Yukardaki test için takip edilecek sıra şöyledir :

1 — Varsayım ifade edilir,

$$\sum_{i=1}^{p} a_i \overline{X}_i$$

Ortalamalar arasındaki ilişki, kesin olarak, test yapılmadan evvel ortaya konmalıdır.

2 — Güvenirlik derecesi seçilir,

3 — (11) eşitliğinden F — istatistiği hesaplanır; kontrol tek taraflıdır.

4 — Seçilen güvenirlik derecesi ve serbestiyet derecelerine göre, F — tablosundan kritik değer bulunur,

5 — Hesaplanan değer, kritik değerden büyükse, varsayım red edilir.

Yukardaki varsayım testinde, F — dağılım tablosunu kullanmak zorunluluğu yoktur : (11) eşitliğinin kare kökü t — dağılımı gösterir; dağılımın serbestiyet derecesi $(\Sigma N_i - P)$ dir.

B — Varyanslarla İlgili Testler

Varyansla ilgili varsayım testlerinde, genelikle, şu durumlar söz konusudur :

Ba. Tek bir varyansla ilgili varsayım testi,

Bb. İki varyansla ilgili varsayım testi,

Bc. İkiden fazla varyansla ilgili varsayım testi.

Ba. Tek bir varyansla ilgili varsayım testi

Bu tür testte, toplum varyansının belirli bir değere eşit olup olmadığı incelenir. Söz konusu toplumdan alınan örnekten hesaplanan örnek varyansının, önceden verilen bir değere eşit olup olmadığını tayinde aşağıdaki şekilde hareket edilir :

1 — Varyansla ilgili varsayım ifade edilir,

$$\sigma^2 = \sigma_0^2$$

2 — Güvenirlik derecesi seçilir,

3 — Varsayımın testinde kullanılacak istatistik

$$K^2 / sd. = S^2 / \sigma_0^2 \tag{13}$$

olup, (13) eşitliği (khi karesi / serbestiyet derecesi) denilen bir dağılım gösterir.

4 — Seçilen güvenirlik derecesi ve (N — 1) serbestiyet derecesine göre, tablodan kritik değerler alınır. (13) eşitliğine göre hesaplanan değer bu kritik değerlerden küçük veya büyükse, varsayım red edilir. Kritik değerlerin bulunmasında (K^2 / Serbestiyet derecesi) özel tablosunun bulunması şart değildir. Zira, Khi karesi tablosundan alınacak değerler, serbestiyet derecesine bölünmekle, Khi karesi / Serbestiyet derecesi tablosunun değerleri elde edilir.

Tek bir varyansla ilgili varsayım testi genellikle tek taraflıdır. Buna göre, bir tek kritik değer, dolayısiyle bir tek kritik bölge bulunacaktır.

Bb. İki varyansla ilgili varsayım testi

İki ayrı toplumdan alınacak örneklerden hesaplanan S_1^2 ve S_2^2 örnek varyansları yardımı ile, bu toplumların varyanslarının eşitliği araştırılabilir. Bu amaçla takip edilecek hareket tarzı şöyledir :

1 — Varyanslarla ilgili varsayım ifade edilir,

$$\sigma_1^2 = \sigma_2^2$$

2 — Güvenirlik derecesi seçilir,

3 — Varsayımın testinde kullanılacak istatistik

$$F = S_1^2 / S_2^2$$

eşitliğinden hesaplanır. (14) eşitliği F — dağılımı gösterir. Dağılımın serbestiyet derecesi $(N_1 - 1)$ ve $(N_2 - 1)$ dir. Varsayımın testi iki taraflıdır.

4 — Seçilen güvenirlik derecesi ve $(N_1 - 1)$, $(N_2 - 1)$ serbestiyet derecelerine göre F — tablosundan kritik değerler alınır.

5 — (14) eşitliğine göre hesaplanan istatistik kritik değerlere göre belirlenen kritik bölgeye düşerse, varsayım red edilir.

Toplumlardan birisinin varyansının diğerinden küçük olduğu şeklindeki bir varsayımın testi ise tek taraflıdır. Bu durumda, tek bir kritik değer ve kritik bölge mevcuttur.

Genellikle, F — tablolarında, bütün güvenirlik dereceleri ile serbestiyet derecelerine tekabül eden değerler gösterilmezler. Böyle tablolarla karşılaşıldığında,

α = Güvenirlik derecesi,

$N_1 - 1$ = Birinci varyansın serbestiyet derecesi,

$N_2 - 1$ = İkinci varyansın serbestiyet derecesi olmak üzere

$$F\ (1 - \alpha),\ (N_1 - 1);\ (N_2 - 1) = \frac{1}{F\ \alpha,\ (N_2 - 1);\ (N_1 - 1)} \qquad (15)$$

eşitliğinden faydalanılır.

Bc. İkiden fazla varyansla ilgili test

Bu tür varsayımların testinde kullanılmak üzere geliştirilmiş birden fazla istatistik vardır. Burada, daha genel durumlara uygulanabilmesi nedeni ile, yalnız Barlett testi örnek olarak verilecektir.

p tane normal dağılımlı toplumdan, p tane N_i üniteden müteşekkil tesadüfi örnek alındığını farz edelim. Örneklerdeki ünite sayılarının birbirlerine eşit olma şartı yoktur. Testin uygulanmasında şu şekilde hareket edilir :

1 — Varsayım ifade edilir,

$$\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \ldots\ldots\ldots = \sigma_p^2$$

Yukardaki normalite ve tesadüfiyet şartları yerine getirilmişse,

$$K^2_{p-1} = \frac{2{,}3026}{C}\left[(N-p)\log S_p^2 - \sum_{i=1}^{p}(N_i - 1)\log S_i^2\right] \qquad (16)$$

istatistiği, yaklaşık olarak khi karesi dağılımı gösterir. Dağılımın serbestiyet derecesi $(p-1)$ dir. (16) eşitliğindeki C katsayısı

$$C = 1 + \frac{1}{3(p-1)}\left[\sum_{i=1}^{p}\frac{1}{N_i-1} - \frac{1}{N-p}\right]$$

$$S_p^2 = \frac{\sum_{i=1}^{p}(N_i-1)\,S_i^2}{N-p}$$

dir. S_i^2 ler ne kadar farklı iseler (16) istatistiği o kadar büyük bulunacaktır.

2 — Güvenirlik derecesi seçilir,

3 — Varsayımın testinde kullanılacak istatistik (16) eşitliğinden bulunur. Varsayımın testi tek taraflıdır.

4 — Seçilen güvenirlik derecesi ve $(p-1)$ serbestiyet derecesine göre, khi karesi tablosundan kritik değer bulunur.

5 — (16) eşitliğine göre hesaplanan değere kritik değerden büyükse varsayım red edilir.

C — Uygunluk Testi

Yukarda açıklanmasına çalışılan iki grup varsayım testlerine ek olarak, örnekten bulunan dağılım ile belirli bir teorik dağılımın uygunluk varsayımı ve bu varsayımın testi zikredilebilir. Böyle bir varsayımın testinde kullanılacak istatistik khi karesidir ve aşağıdaki şekilde hareket edilir :

1 — Teşkil edilen k tane kademeye, giren örnek ünitesi sayısı (veya değeri) bulunur.

2 — Bu kademelerde, kabul edilen dağılıma göre, bulunması gereken teorik ünite sayıları tespit edilir.

3 — Mütekabil kademelerde, örnek ünitesi sayı ile teorik ünite sayısı arasındaki farkın karesi, kademedeki teorik örnek ünite sayısına bölünür. Bu işlem, bütün kademeler için ayrı ayrı yapılır.

4 — Her kademe için hesaplanan bu değerler toplanır. Bu şekilde bulunan istatistik, $(k-1-p)$ serbestiyet derecesine sahip khi karesi dağılımı gösterir. Serbestiyet derecesi ifadesindeki (p) terimi, dağılımı karakterize eden parametre değerleri sayısıdır. Meselâ, normal dağılım halinde, bilinmesi gereken parametre sayısı, ortalama ve varyans olmak üzere, iki tane, yani $p = 2$ dir.

5 — Seçilen güvenirlik derecesi ve $(k-1-p)$ serbestiyet derecesine göre, khi karesi tablosundan kritik değer bulunur. Varsayımın testi tek taraflıdır. 4. şıkta hesaplanan değer kritik değerden büyükse, varsayım red edilir.

Testlerle ilgili tartışmalar

Yukarıda, istatistiksel yargılarda en çok kullanılan testler ana hatları ile açıklanmaya çalışılmıştır.

Herahangi bir varsayımın testinde kullanılan istatistikler yardımı ile güvenilir yargılara varılabilmesi, kaba hataların dışında, söz konusu istatistiğe esas olan kabullerin gerçekleşmesine bağlıdır. En kuvvetli istatistiksel testler, kapsamı geniş kabullere istinat eden istatistikleri kullananlardır. Meselâ t veya F istatistiklerini kullanan testler bu türdendirler. Bu istatistiklerin istinat ettiği esaslar gerçekleştiği taktirde, yanlış bir varsayımı red etme ihtimali, yani testin kuvveti en fazladır.

Yapılan açıklamalardan da görüleceği gibi, meselâ t testinin uygulanabilmesi, aşağıdaki şartların gerçekleştirilmesini gerektirir :

1 — Toplumdan alınan ünitelerin seçimi tesadüfi ve bağımsız olmalıdır; diğer bir deyişle; örnek üniteleri her türlü kişisel tercihlerin dışında seçilmeli ve herhangi bir örnek ünitesinin seçimi bir diğerinin örneğe dahil olmasını engellememeli veya ihtimalini azaltmamalıdır.

2 — Örnek, normal dağılımlı toplumdan alınmış olmalı veya örnekteki ünite sayısı yeter üyüklükte, meselâ otuz veya daha çok olmalıdır.

3 — Örnek ünitelerine ilişkin nitelikler ölçülebilmeli veya rakamla ifade edilebilmeli, böylece bunlar üzerinde matematik işlemleri yapılabilmelidir.

F — dağılımı halinde, yukarıdakilerine ek olarak, toplumların varyansları eşit ve toplanabilir nitelikte — additive — yani toplumları etkileyen faktörlerin doğrusal bir fonksiyonu olmalıdır.

Uygulamada, eşit varyanslılık dışında, diğer kabullerin gerçekleşip gerçekleşmediği pek incelenmemektedir. — Gerçi bu amaçla geliştirilmiş istatistikler mevcuttur —. Kabullerin gerçekleşip gerçekleşmediğinin önceden bilinmemesi, varılacak yargının istenilen güvenirlik sınırı içinde kalıp kalmadığı şüphesini uyandıracaktır. Bununla beraber, sözü edilen kabullerden «küçük» ayrılmaların yargılarımıza olan olumsuz etkisini örnekteki ünite sayısını arttırmakla, ihmal edilebilir bir seviyeye düşürebileceği gerek teorik ve gerekse uygulamalarla gösterilmiştir.

Yukarıda genel hatları ile açıklanmaya çalışılan testler ve bunların hangi varsayımların kontrolunda kulanılacağı Şema 1'de gösterilmiştir. Şemadaki kanallar takip edilmek suretiyle varsayım çeşidine uygun test türünü kararlaştırmak mümkündür.

FAYDALANILAN ESERLER

Anderson R.L., Bancroft T.A., 1952, Statistical Theory in Research McGraw - Hill N.Y.

Dixon W.J., F.J. Massey Jr. 1956, Introduction to Statistical Analysis, McGraw - Hill, N.Y.

Fisher R., 1966, The Design of Experiments, London.

Fraser D.A.S., 1964, Statistics, An Introduction, John - Wiley and Sons, N.Y.

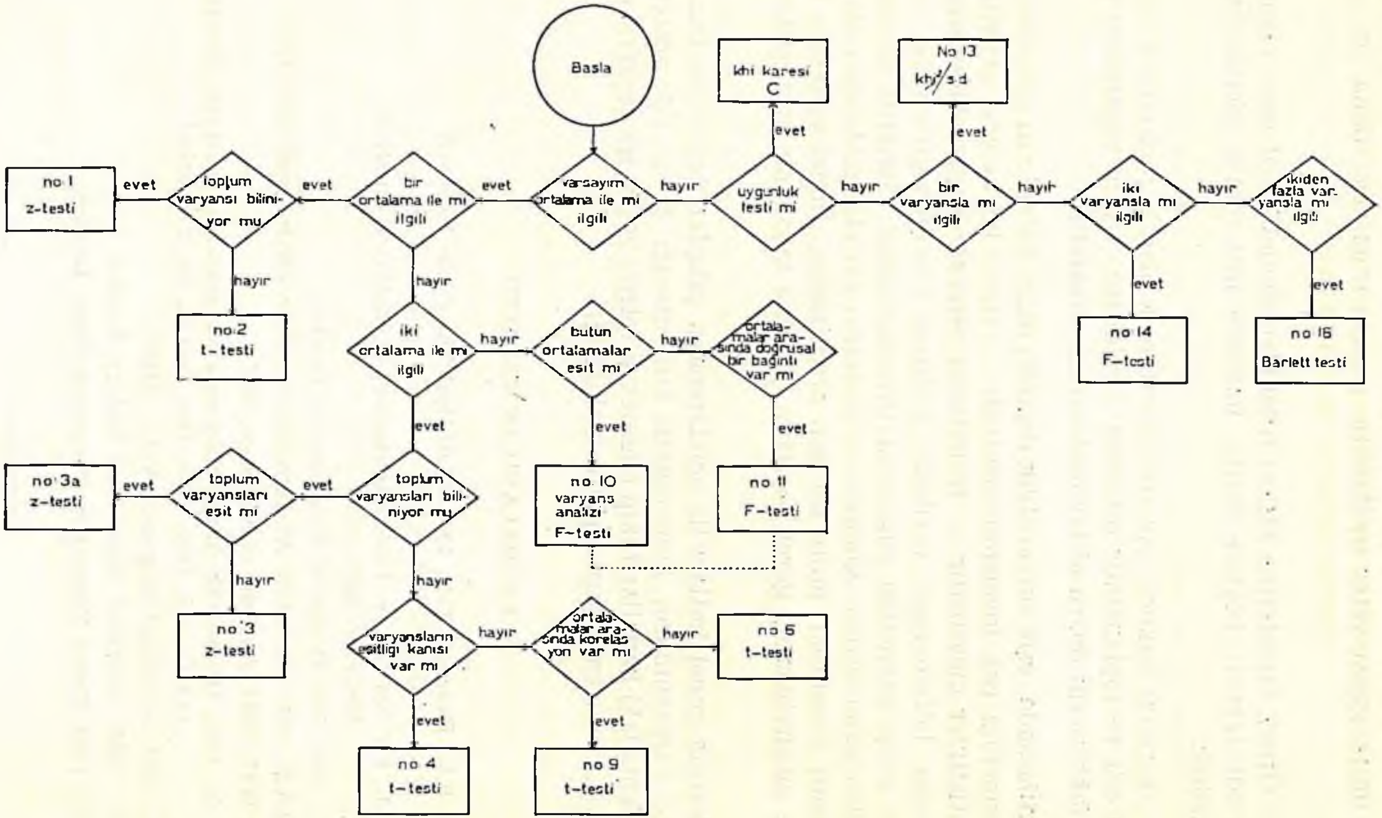
Gunther W.C., 1964, Analysis of Variance, N.Y.

Kalıbsız A., 1959, Ormancılık Araştırmalarında Matematik - İstatistik Metodlarının Önemi. Orm. Fak. Der. Seri B, Sa. 2 İstanbul.

Li J.C.R., 1964, Statistical Inference, Vol. I. Mich.

Mather K., 1966, Statistical Analysis in Biology, London.

Prodan M., 1966, Forest Biometrics, Pergoman Press, London.



Sema I Hangi testin hangi durumlarda kullanılacağını gostermektedir