σ Journal of Engineering and Natural Sciences
Mühendislik ve Fen Bilimleri Dergisi

Sigma
2005/3

SOME TOPICS TO BE TAKEN UP IN STATISTICS

Mehmet GENCELİ[*]

Yıldız Teknik Üniversitesi,Fen-Edebiyat Fakültesi, İstatistik Bölümü, Davutpaşa - İSTANBUL

Geliş/Received: 01.02.2005 Kabul/Accepted: 05.09.2005

ABSTRACT

This expository article discusses the elementary teaching of one-sided confidence intervals along with one and two sided prediction intervals.
For being a complementary part of one-sided hypothesis test one-sided confidence intervals should be considered as an indispensable tool in Statistics. Despite this fact Turkish Statistics Literature has only focused on two-sided confidence intervals.On the other hand, prediction intervals for the univariate data have not been taken up to date. Basic Turkish Statistics textbooks invariably introduce the topic of prediction in the context of simple regression for the bivariate normal model.Adoption of the Hahn's approach may give prediction the prominent position in basic statistical education that it deserves.
Keywords: ne- and two- sided hypothesis testing, one- and two- sided confidence- and prediction intervals for univariate data, Hahn's approach.
MSC number/numarası: 65C60

İSTATİSTİK ÖĞRETİMİNDE İRDELENMESİ GEREKLİ KİMİ KONULAR

ÖZET

Hipotez testleri ve güven sınırları her düzeydeki İstatistik öğretiminde yadsınmaz bir öneme sahiptir. Buna rağmen Türkçe yazının hemen hemen tümünde hipotez testlerinin tek veya çift taraflı yapılmış olmasına bakılmaksızın, H_0 reddi durumunda çift yönlü güven sınırları oluşturulmaktadır. Böylece yönün bilinmesinden ötürü kazanılmış bilgi kaybedilerek α ile alınan kararın sonucu ($\alpha / 2$) ile uygulanmaktadır. İkinci bir olgu ise Türkçe yazında tek değişkenli dağılımlar için öngörü sınırlarına günümüze kadar hiç yer verilmemiş olmasıdır. Bundan ötürü adı geçen sınırlar tanıtılarak almaşık hesaplama yöntemleri açıklanmıştır.
Anahtar Sözcükler: Tek ve çift taraflı hipotez testleri, Tek ve çift taraflı güven ve öngörü sınırları Hahn tabloları.

1. GİRİŞ

1950'li yılların başında İstatistik dersleri bir elin parmaklarını geçmeyen yüksek öğrenim kurumlarında gene birkaç öğretim üyesi tarafından veriliyordu. O dönemlerde İstatistiğin bağıl olarak daha yoğun okutulduğu ve günümüzdeki yan dala karşılık gelen İstatistik Disiplinin'den mezun veren İ.Ü. İktisat Fakültesin'de Matematik dersi bile bulunmuyordu.

Ancak 1950'li yılların ortalarından itibaren İstanbul'da Prof.Dr. Haydar Furgaç, Prof.Dr. Fazıl Gülçur, Ankara'da da Prof.Dr.Necati İşçil ve Prof.Dr.Orhan Düzgüneş o zamanki adıyla Sondaj, şimdiki adıyla Örneklemeyi tanıtarak İstatistiği betimsellikten kurtararak Hipotez Testleri ve Güven Sınırları'nın odak olduğu Tümevarım İstatistiğe doğru yönlendirdiler.

[*] e-posta: mehmetgenceli@yahoo.com, tel: (0212) 449 17 09

Kuşkusuz Hacettepe Üniversitesi İstatistik Bölümü'nün 1967 yılında kurulması ile İstatistik öğretiminin araçtan amaç haline dönüşmesi, YÖK Yasa'sı ile oluşturulan İstatistik ve/veya Ekonometri Bölümleri'nin kurulması, en önemlisi de bilgisayar teknolojisi ve programlarındaki gelişmeler kanımca İstatistik öğretiminin ivme kazanmasındaki başlıca nedenler olarak sayılabilir.

Farklı yoğunlukta da olsa halen tüm Üniversitelerde İstatistik öğretimi yapılmakta, öğrenci kitlesi de Mühendislik'ten İşletme ve İktisat'a, Sosyoloji'den Eğitim Bilimlerine kadar uzanmaktadır. Bunlara sayısı 15-16 olan İstatistik Bölümlerin'deki İstatistik amaçlı öğretim de eklenirse öğretimdeki heterojenlik kendiliğinden ortaya çıkmaktadır.

Tüm bu çeşitliliğe rağmen, İstatistiğin günümüzdeki tanımının "verilerden bilgi edinerek belirsizlik altında karar almaya yarayan yöntemler topluluğu" [1] olduğu göz önüne alınırsa hipotez testleri ve güven sınırlarının İstatistik öğretiminde yadsınmaz bir yeri ortaya çıkmaktadır.

Bu nedenle de yazıda hipotez testleri ve güven sınırları öğretimindeki boşlukların bir dereceye kadar doldurulması ve günümüzde güven sınırlarının tamamlayıcısı olarak İstatistik yazınında yer almaya başlayan öngörü sınırlarının irdelenmesi amaçlanmıştır.

2. HİPOTEZ TESTLERİ VE GÜVEN SINIRLARI

Türkçe İstatistik yazınının hemen hemen tümünde herhangi bir θ parametresi için $H_0 : \theta = 0$, $H_1 : \theta \neq 0$ veya $H_0 : \theta = \bar{\theta}$, $H_1 : \theta \neq \bar{\theta}$ gibi çift veya $H_0 : \theta \leq 0$, $H_1 : \theta > 0$; $H_0 : \theta \geq 0$, $H_1 : \theta < 0$ veya $H_0 : \theta \leq \bar{\theta}$, $H_1 : \theta > \bar{\theta}$; $H_0 : \theta \geq \bar{\theta}$, $H_1 : \theta < \bar{\theta}$ gibi tek taraflı hipotez testlerine yer verilmekte ve ana kütle normal dağılmakta ise, $X \sim N(\mu, \sigma_X^2)$ örnek için de normal dağılım kullanılarak veya ana kütle normal dağılmıyorsa $n > 30$ için $\hat{\theta} \sim N(\theta, \sigma_{\hat{\theta}}^2)$ varsayılarak hüküm verilmektedir.

Hipotez testi sonucunda da bana göre yanlış bir ifade olan "H_0 kabul " denilmekte veya H_0 reddedilmektedir.

H_0 zaten önceden kabul edilmiş bir hipotez olduğuna göre "H_0'ın reddi için yeterli kanıt veya herhangi bir neden olmadığı" ifadesi sanırım daha doğru bir ifadedir. Buna karşın, H_0'ın reddi halinde doğru ifade kullanılmakta, fakat hipotez testinin tek veya çift taraflı yapıldığına bakılmaksızın, H_0'ın reddi halinde parametre için

$$\hat{\theta} - z(\alpha/2)\sigma_{\hat{\theta}} \leq \theta \leq \hat{\theta} + z(\alpha/2)\sigma_{\hat{\theta}} \quad \text{veya} \tag{1}$$

$$\hat{\theta} - t(n-1;\alpha/2)s_{\hat{\theta}} \leq \theta \leq \hat{\theta} + t(n-1;\alpha/2)s_{\hat{\theta}} \tag{2}$$

güven sınırları yazılmaktadır.

Hipotez testlerinin tek taraflı yapılması halinde önemli bir hata yapılmaktadır. Bu da hipotez testine ilişkin kararın α ile alınmasına karşın güven sınırlarının ($\alpha/2$) ile saptanmasıdır. Başka bir deyişle, hipotez testine ilişkin kararın α veya ($\alpha/2$)'ye göre alınmasına bakılmaksızın güven sınırları daima ($\alpha/2$) ile hesaplanmaktadır.

Halbuki hipotez testi çift taraflı ise güven sınırları çift taraflı, hipotez testinin tek taraflı olması halinde güven sınırı da tek taraflı olmak zorundadır.

Tek taraflı güven sınırları sadece Türkçe yazında değil aynı zamanda İngilizce yazında da sorundur. Giriş düzeyindeki çoğu İngilizce İstatistik kitaplarının tek taraflı güven sınırını işlememesi eleştiri konusu olmaktadır [2]. Nitekim İngilizce yazında yapılan bu hata fark edilerek İstatistik kitaplarında son yıllarda tek taraflı güven sınırlarına yer verilmeye başlanılmıştır [3].

Bu bağlamda tek taraflı güven sınırı, $H_0 : \theta \geq 0$ veya $H_0 : \theta \geq \overline{\theta}$ için, H_0'ın reddi halinde, $z = [(\hat{\theta} - \theta) / \sigma_{\hat{\theta}}]$ 'dan hareketle $\%100\gamma = \%(1-\alpha)$ olasılıkla güven sınırının alt sınırı, yani $\%100\gamma$ olasılıkla

$\hat{\theta} - z(\alpha)\sigma_{\hat{\theta}}$ veya (3)

$\hat{\theta} - t(n-1;\alpha)s_{\hat{\theta}}$ (4)

Şeklinde parametrenin daha küçük değerde olamayacağı elde edilir.

$H_0 = \theta \leq 0$ veya $H_0 : \theta \leq \overline{\theta}$, $H_1 : \theta > 0$ veya $H_1 : \theta > \overline{\theta}$ için ise bu kez $-z = [(\hat{\theta} - \theta) / \sigma_{\hat{\theta}}]$ 'dan güven sınırı

$\hat{\theta} + z(\alpha)\sigma_{\hat{\theta}}$ veya (5)

$\hat{\theta} + t(n-1;\alpha)s_{\hat{\theta}}$ (6)

üst sınırına ulaşılır.

Çift taraflı güven sınırları ile karşılaştırıldığında tek taraflı güven sınırı parametre tahmincisi veya diğer adıyla istatistiğe daha yakın başlamakta bu da parametre tahmin edicilerinde aranılan özelliklerden biri olan kesinliği arttırmaktadır.

Sözgelimi $X \sim N(\mu, \sigma^2)$ bir ana kütleden çekilen 18 birimlik bir örneğin ortalaması $\overline{X} = 39$, ortalamanın standart hatası da 6 ise, $s_{\overline{x}} = 6$, μ için %95 olasılıkla güven sınırları

$39 - (2.11).6 < \mu < 39 + (2.11).6$

$26.34 < \mu < 51.66$

olacaktır. Bu bağlamda gene %95 olasılıkla tek taraflı güven sınırının alt hududu

$39(1.74).6 = 39 - 10.4 = 28.56$

çıkacaktır. Buna göre %95 olasılıkla ana kütle ortalamasının en az 28.56 olduğu söylenebilecektir. Öte yandan hesaplanmış olan 28.56, çift taraflı güven sınırlarının alt hududu olan 26.34'e kıyasla, örnek ortalaması 39'a daha yakındır.

Bu gibi tek taraflı hipotez testi, dolayısıyla tek taraflı güven sınırı için şu örnek verilebilir [4]:

Uğraşı alanı nedeniyle yüksek seyahat giderleriyle karşı karşıya gelen bir şirket ertesi yıl için bütçesinde ödenek ayırmak istemektedir. Geçmişte yapılan 83 seyahatin kayıtlarından seyahat giderleri ortalaması 1286$ ve ortalamanın standart hatası $\sigma_{\overline{x}} = 71.03$ hesaplanmıştır. %95 olasılıkla ortalama için güven sınırları

$1286 - (1.96)(71.03) < \mu < 1286 + (1.96)(71.03)$

$1147 < \mu < 1425$

olmasına rağmen şirket açısından seyahat giderlerinin üst hududu çok daha önemlidir. Bunun için de gene %95 ile

$\overline{X} + z(\alpha)\sigma_{\overline{x}} = 1286 + (1.645)(71.03) = 1403\$$

üst hudut hesaplanmıştır. Böylece %5 hata ile ortalamanın 1403$’ dan daha fazla olamayacağını söylemek daha doğrudur.

Kısaca vurgulamak gerekirse, ilişkinin yönü belli olup tek taraflı hipotez testinin yapıldığı durumlarda, H_0 reddedildiği takdirde tek taraflı güven sınırının uygulanması çok daha doğru olacaktır.

3. ÖNGÖRÜ SINIRLARI

Örneklemenin uygulandığı birçok olay için ortalama ve dağılma ölçülerinin yanında birimlere ilişkin bilgi alınmak istenilebilir. Böyle durumlarda birim hakkında bilgi veren öngörü ve tolerans sınırları ile parametrik olmayan sınırlar söz konusu olmaktadır. Bu sınırlar arasında öngörü sınırları ön planda olduğu gibi Mühendislik’te önemli bir konuma sahip tolerans sınırlarını diğer iki sınırdan farklı bir yere koymak gerekir.

Öngörü sınırları Türkçe yazında regresyon konusunu işleyen kitapların hemen hemen tümünde uzunca bir süredir yer almasına [5] rağmen, tek değişkenli veriler için bu kavrama Türkçe yazında ne yazık ki henüz rastlanılmamıştır.

Normal dağılan bir ana kütleden , $Y \sim N(\mu, \sigma_Y^2)$ önceden çekilmiş olan n birimlik örnekten elde edilen bilgiye dayanarak aynı ana kütleden sonradan çekilen k birimi, k=1,2,.... belirli bir olasılıkla kapsayan aralık öngörü sınırları olarak adlandırılmaktadır.

Öngörü sınırlarının önemini vurgulamak açısından çeşitli örnekler verilebilir. Herhangi bir araç/ gereç tamiri için yedek parça bekleyen bir servis için parçanın ortalama geliş süresinden ziyade en geç ne zaman geleceği önemlidir. Kalp pili takılacak bir hasta için pilin ortalama dayanma süresi değil, en az dayanma süresi hayati önemdedir.

Sayılan bu örneklerden birimlerden alınacak bilginin de ortalamadan alınacak bilgi kadar önemli olduğu, hatta bazen onun önüne geçtiği savunulabilir.

Anılan öngörü sınırlarının ne kadar önemli olduğu İngilizce yazında daha 1969’da vurgulanmasına [6] rağmen basit rastlantısal örnekleme kapsamındaki öngörü sınırları bu yazında dahi hak ettiği ilgi ve yeri ancak son yıllarda bularak [7] İstatistik kitaplarına girmiştir. [8]

Bir ana kütleden çekilen tek birim için $\%100\gamma = \%100(1-\alpha)$ öngörü sınırları

$$\mu \pm z(\alpha/2)\sigma$$

şeklindedir. Ancak uygulamada μ ve σ bilinmediğinden normal dağılan bir ana kütleden çekildiği varsayılan, ortalaması $\bar{Y} = \sum Y_i / n$, standart sapması da $s = [\sum (Y_i - \bar{Y})^2 / (n-1)]^{1/2}$, i=1,2,....,n olan n birimlik örnekten hareket etmek çok daha gerçekçidir.

Öngörü hesabı için önerilen iki yanlı yaklaşımdan bir Scheffe tarafından önerilen, Varyans Analizindeki kontrastlara ve F dağılımına dayanan [9] yaklaşım, diğeri ise Bonferroni eşitsizliğini ve (n-1) serbestlik dereceli t- dağılımını kullanan yaklaşımdır [10].

Birinci yaklaşımda ortalaması $\bar{Y}$, standart sapması s olan n birimlik bir örnekten sonradan çekilen k birim için öngörü sınırları

$$\bar{Y} \pm [F(k, n-1; \gamma]^{1/2} (n+1/n)^{1/2} s \qquad (7)$$

olarak bulunurken ikinci yaklaşımda da bu sınırlar

$$\bar{Y} \pm t[n-1; 1-(1-\gamma)/2k](n+1/n)^{1/2} s \qquad (8)$$

şeklindedir.

Uygulamada ikinci yaklaşım tercih edildiğinden burada da ikinci yaklaşım tercih edilmiştir. Bununla beraber k=1 için $t^2[n-1;1-(1-\gamma)/2] = F(1,n-1;\gamma]$ eşitliği söz konusu olduğundan her iki yaklaşım da aynı sonucu verecektir.

Basitliğinden dolayı, önce, sonradan çekilen tek birim , sonra da sonradan çekilen birden çok birim irdelenmeye çalışılacaktır.

Bir ana kütleden n birimlik bir örnek çekilip $\bar{Y}$ ve s hesaplandıktan sonra aynı ana kütleden bu kez X ile öngörülen , $\bar{Y}$ ve s'den bağımsız bir birimin sonradan çekildiği varsayılırsa

a- $\varepsilon = (X-\bar{Y})$ öngörü hatası olup $E(X-\bar{Y}) = E(\varepsilon) = 0$ dır.

b- $V\text{ar}(X-\bar{Y}) = VarY + Var\bar{Y} = \sigma^2 + \sigma^2/n = \sigma^2[(n+1)/n]$ olmaktadır [11]. Birinci varyans birimlerin Y_i, i=1,2,...,n ikinci varyans ise, $\sigma^2_{\bar{Y}}$, örnek ortalamalarının μ etrafındaki dağılımının ölçüsüdür. Zaten $\sigma_{\bar{Y}}$ ortalamanın standart hatasıdır.

μ bilinmeyip her parametre yerine örnek ortalaması ikame edildiğinden öngörü varyansı için de

$$s^2(X-\bar{Y}) = s[(n+1)/n]^{1/2} \tag{9}$$

kullanılmalıdır.

c- X ve $\bar{Y}$ birbirinden bağımsız normal değişkenler olduğundan bu iki normal değişkenin birbirinden farkını veren $Z = (X-\bar{Y})$ değişkeni de normal dağılacaktır: $Z{\sim}N(0,\sigma^2_{X-\bar{Y}})$

d- $[(n-1)s^2/\sigma^2]$ ise Z 'den bağımsız olarak (n-1) serbestlik dereceli bir χ^2 dağılımını gerçeklemektedir.

$z = Z/[\sigma((n+1)/n)]^{1/2}$ yazılırsa $z{\sim}N(0,1)$ standart normal dağılıma dönüşecektir.

Diğer taraftan $u = \left\{[(n-1)s^2/\sigma^2]1/(n-1)\,]\right\}^{1/2} = (s/\sigma)$ (11)

yazılarak (z/u) oranı oluşturulursa, bu oran $s.d = (n-1)$ serbestlik derecesindeki t dağılımını verecektir. Böylece k=1 için $\%100\gamma = \%100(1-\alpha)$ çift taraflı öngörü sınırları

$$\bar{Y} \pm t[n-1;1-\alpha/2](n+1/n)^{1/2}s \tag{10}$$

olacaktır. $(1+\gamma)/2 = 1-\alpha/2$ eşitliğinden hareketle aynı öngörü sınırları bu kez

$$\bar{Y} \pm t[n-1;(1+\gamma)/2](n+1/n)^{1/2}s \tag{11}$$

şeklinde yazılabilir [11].

(10) ve (11) formülleri aslında

$$\bar{Y} \pm t[n-1;1-\alpha/2][(1/m)+(1/n)]^{1/2}s \tag{12}$$

formülünün m=1 için özel halidir. m ise çekilen örnek sayısını göstermektedir. Genelde her dönem için sadece bir örnek çekildiğinden m=1 için (10) formülü (8) formülüne dönüşmekte ve bu biçimde kullanılmaktadır.

Güven sınırlarında olduğu gibi hiç kuşkusuz öngörü sınırları için de tek taraflı sınırlar oluşturulabilir:

k=1 için, yani sonradan çekilen tek bir birim için %(1- α) olasılıkla alt hudut

$$\overline{Y} - t[n-1;\alpha](n+1/n)^{1/2} s \quad (13)$$

üst hudut ta

$$\overline{Y} + t[n-1;\alpha](n+1/n)^{1/2} s \quad (14)$$

formüllerine göre belirlenmektedir [11].

Ayrıca X'in , yani sonradan çekilen birimin örneğin ana kütlesinden çekilip çekilmediği de sınanabilir[12]:

H_0: X ve örnek birimleri normal dağılan aynı ana kütleden çekilmiştir.

H_1: X ve örnek birimleri ayrı ana kütlelerden çekilmişlerdir.

Hipotez testine ilişkin test istatistiği de

$$t = (X - \overline{Y})/(n+1/n)^{1/2} s \quad (15)$$

olup $t(n-1;1-\alpha/2)$ kritik değerler ile karşılaştırılıp hüküm verilecektir.

Aynı zamanda (15) yardımıyla tek taraflı hipotez testi de yapılabilir. Ancak tek taraflı yapıldığında X'in sonradan çekilen biriminin ortalamadan büyüklüğü veya küçüklüğü sınanmaktadır. Kritik değer ise $t(n-1;1-\alpha)$ dır.

k=1 için aşağıdaki uygulama verilebilir:

Uygulama 1

Yaşamı böbrek nakline bağlı olan bir hastaya en geç 5 gün içinde nakil gereklidir. Geliş süreçlerinin normal dağıldığı böbreklerden 8 birimlik bir örnek çekilerek gün olarak geliş süreleri saptanmıştır.

Y:10,9,7,10,3,9,12,5 $\overline{Y}$ =8.125 s =2.94897 $s_{\overline{x}} = 1.0426183$ X =5

Ortalama için %95 güven sınırları

$\overline{Y} \pm t[7;0.975]s_{\overline{x}}$ $8.125 \pm (2.365)(1.0426183)$

$$5.66 < \mu_Y < 10.59$$

gün olduğundan %5 hata ile hastanın hayatta kalma şansı bulunmamaktadır.

Ancak "en geçin" karşılığı olarak çift taraflı güven sınırı yerine gene %95 için oluşturulan

$$\overline{Y} + t[7;0.95] = 8.125 + (1.895)(1.0426183) = 10.10$$

tek taraflı güven sınırına göre %95 olasılıkla böbreğin ortalama geliş süresi 10.10 günden geç değildir.

Ortalama için tek taraflı güven sınırı uygulandığı takdirde durum farklıdır. %5 hata payı ile böbreğin ortalama olarak 1 ila 10 gün arasında geleceği söylenebilecektir. Çift taraflı güven sınırlarında alt hudut 5.66 olmasında dolayı hastanın yaşama şansı bulunmaz iken, burada böbreğin 6 günden önce gelişine bağlı olarak hastanın yaşama şansı vardır. Bununla beraber ortalama için güven sınırları birimlerin ortalama etrafındaki dağılmasını dikkate almadığından hayati önemdeki bu olayda her türlü belirsizliğin göz önüne alınması gerekir. Dolayısıyla da çözüm için güven sınırları yerine öngörü sınırlarının kullanılması daha yerinde bir karardır.

Bunun için de (8) veya (9) formülü ile

$$8.125 \pm (2.365)(2.94897)(1.0606)$$

$$0.728 \le X \le 15.52$$

Böylece böbreğin geliş sürecinin; %95 olasılıkla 0.7 ila 15.5 gün arasında olacağı ifade edilebilecektir. Öngörü alt hududu da 0.7 gün olması nedeniyle güven sınırlarının aksine hastanın yaşama şansı bulunmaktadır. Diğer taraftan böbreğin gene %95 olasılıkla en geç kaç günde geleceği sorusunun yanıtı da (12) formülünden yararlanılarak

$8.125 + (1.895)(2.94897)(1.0606) = 14.05$ gün bulunacaktır.

Problem için hipotez testi de yapılabilir:

H_0: X =5 ve örnek aynı ana kütleden çekilmiştir.

H_1= X =5 ve örnek farklı ana kütlelere ilişkindirler.

(15) formülü uygulanırsa t=(5-8.125)/2.94897=-1.06 elde edilebilir. Test istatistiği t[7;0.975]=-2.365' den büyük olduğundan H_0 reddi için yeterli bir neden bulunmamaktadır. X, örneğin çekildiği ana kütleden çekilmiştir.

k>1 için ise gene $Z_i = X_i - \bar{Y}$ çıkış noktasıdır. $Z_1, Z_2, ..., Z_k$ ortalamaların 0, varyansların da $\sigma^2 (n + 1/n)$ olduğu çok değişkenli bir normal dağılıma uymaktadır. Buradan da k=1'e koşut olarak oluşturulan

$$t_i = (X_i - \bar{Y}) / [(n + 1/n)^{1/2} s] \qquad \text{i=1,2,.........,k} \qquad (16)$$

değişkenlerinin bileşik dağılımı bu kez çok değişkenli t- dağılımına uymaktadır.

Ancak k>1 için çözüm k=1 kadar kolay değildir. Örneğin k=1, $\gamma = 0.95$ için $[1 - (1 - \gamma)/2] = [1 - (1 - 0.95)/2] = \%97.5$ için t- çizelge değeri mevcut olmasına rağmen, k=2 $\gamma = 0.95$ karşılığı $[1 - (1 - \gamma)/2k] = [1 - 0.05/4] = \%98.75$ veya k=3 $\gamma = 0.99$ için $[1 - (1 - 0.99)/6] = \%99.83$ t çizelgesinde bulunmayan değerlerdir. Çok değişkenli t-dağılımının (n-1) serbestlik derecesi yanında $k \times k$ korrelasyon matrisini de kapsaması nedeniyle çoğu değerlerin t çizelgesinde yer almaması hiç şaşırtıcı değildir. Tek ayrıcalık ise Dunn tarafından sadece $\gamma = 0.95$ için hazırlanmış olan çizelgedir.

Çizelge 1. $(1 - \alpha / 2k) = (1 - 0.05 / 2k)$ için t çizelgesi

n \ k	1	2	3	4	5	6	7	8	9	10	15	20	50
5	2.57	3.16	3.54	3.81	4.04	4.22	4.38	4.53	4.66	4.78	5.25	5.60	6.87
10	2.23	2.64	2.87	3.04	3.17	3.28	3.37	3.45	3.52	3.58	3.83	4.01	4.59
15	2.13	2.49	2.70	2.84	2.95	3.04	3.11	3.18	3.24	3.29	3.48	3.62	4.08
20	2.09	2.42	2.61	2.75	2.85	2.93	3.00	3.06	3.11	3.16	3.33	3.46	3.85
24	2.07	2.39	2.58	2.70	2.80	2.88	2.94	3.00	3.05	3.09	3.26	3.38	3.75
30	2.04	2.36	2.54	2.66	2.75	2.83	2.89	2.94	2.99	3.03	3.19	3.30	3.65
40	2.02	2.33	2.50	2.62	2.70	2.78	2.84	2.89	2.93	2.97	3.12	3.23	3.55
60	2.00	2.30	2.47	2.58	2.66	2.73	2.79	2.84	2.88	2.92	3.06	3.16	3.46
120	1.98	2.27	2.43	2.54	2.62	2.68	2.74	2.79	2.83	2.86	3.00	3.09	3.38
∞	1.96	2.24	2.40	2.50	2.58	2.64	2.69	2.74	2.78	2.81	2.94	3.03	3.29

Kaynak: Dunn Olive Jean, "Confidence İntervals For The Means Of Dependent, Normally Distributed Variables, Journal Of American Statistical Association, 54(1959),s.615.

Dolayısıyla çok değişkenli t- çizelgesinde bulunmayan değerler için interpolasyon yapılması ve bu amaçla da çift taraflı öngörü sınırları için

$F_1(u) = \int_{-u}^{+u} .... \int_{-u}^{+u} f(t_1, t_2, ..., t_k; n - 1, \Sigma) dt_1 ......dt_k = \gamma$ integralinin alt ve üst sınırları u'nun belirlenmesi gerekmektedir. $f(t_1, t_2, ..., t_k)$, n-1 serbestlik derecesine ve Σ korrelasyon matrisine sahip t yoğunluk fonksiyonudur. Çözüm sonucunda da

$$P[\bar{Y} - u(n+1/n)^{1/2} s < X_1 \ldots X_k < \bar{Y} + u(n+1/n)^{1/2} s] = \gamma \qquad (17)$$

öngörü sınırları elde edilecektir.

Tek taraflı güven sınırları için de aynı biçimde hareket edilmekle birlikte integral

$$F_2(u') = \int_{-\infty}^{u'} \ldots \int_{-\infty}^{u'} f(t_1, t_2, \ldots, t_k; n-1, \Sigma) dt_1 \ldots dt_k = \gamma \quad \text{olacaktır.} \qquad (18)$$

Görüldüğü gibi gerek t-çizelgelerinin yetersizliği gerekse integral hesapları k>1 için öngörü sınırlarının kolayca saptanmasını olanaksız hale getirmektedir.

4. HAHN ÇİZELGELERİ

Hahn bu olgudan hareketle Bonferroni eşitsizliğine dayanan ikinci yaklaşımı kullanarak önce çift taraflı öngörü sınırları [13], sonra da tek taraflı öngörü sınırı [14] için çizelgeler hesaplamıştır.

Çizelge 2. Çift Taraflı %90 Öngörü Sınırları için k ve n Bileşimleri

Çekilen ek birim sayısı (k)

Önceden çekilen örnek birim mevcudu (n)	*1*	*2*	*3*	*4*	*5*	*6*	*8*	*10*	*12*	*15*	*20*
					90% Öngörü sınırları						
4	2.63	3.33	3.74	4.03	4.26	4.43	4.71	4.93	5.09	5.30	5.55
5	2.34	2.91	3.25	3.48	3.67	3.81	4.04	4.22	4.36	4.53	4.74
6	2.18	2.69	2.98	3.19	3.35	3.48	3.68	3.84	3.96	4.11	4.30
7	2.08	2.55	2.82	3.01	3.15	3.27	3.46	3.60	3.71	3.85	4.02
8	2.01	2.45	2.71	2.88	3.02	3.13	3.30	3.44	3.54	3.67	3.83
9	1.96	2.38	2.63	2.79	2.92	3.03	3.19	3.32	3.42	3.54	3.70
10	1.92	2.33	2.56	2.73	2.83	2.95	3.11	3.23	3.33	3.44	3.59
11	1.89	2.29	2.52	2.67	2.79	2.89	3.04	3.16	3.25	3.36	3.51
12	1.87	2.26	2.48	2.63	2.75	2.84	2.99	3.10	3.19	3.30	3.44
15	1.82	2.19	2.39	2.54	2.65	2.74	2.87	2.98	3.06	3.17	3.30
20	1.77	2.12	2.32	2.45	2.56	2.64	2.77	2.87	2.95	3.04	3.16
25	1.74	2.09	2.27	2.40	2.50	2.58	2.71	2.80	2.88	2.97	3.08
30	1.73	2.06	2.25	2.37	2.47	2.55	2.67	2.76	2.83	2.92	3.03
40	1.71	2.03	2.21	2.34	2.43	2.50	2.62	2.71	2.78	2.86	2.97
60	1.69	2.00	2.18	2.30	2.39	2.46	2.57	2.66	2.72	2.81	2.91
∞	1.64	1.95	2.11	2.23	2.31	2.38	2.48	2.56	2.62	2.70	2.79

Çekilen ek birim sayısı (k)

	1	2	3	4	5	6	7	8	9	10	11	12	15	20
						95% Öngörü sınırları								
4	3.56	4.41	4.92	5.29	5.56	5.79	5.98	6.14	6.28	6.41	6.52	6.62	6.88	7.21
5	3.04	3.70	4.09	4.36	4.58	4.75	4.90	5.02	5.13	5.23	5.32	5.39	5.60	5.85
						99% Öngörü sınırları								
4	6.53	7.94	8.80	9.41	9.88	10.27	10.59	10.87	11.12	11.33	11.52	11.70	12.14	12.70
5	5.04	5.97	6.54	6.94	7.25	7.51	7.72	7.91	8.07	8.22	8.35	8.47	8.77	9.15

Kaynak: Hahn Gerald J., "Additional Factors For Calculating Prediction İntervals For Samples Form A Normal Distribution", Jasa 65 (1970),1669

Çizelge 3. Çift Taraflı %95 Öngörü Sınırları için k ve n Bileşimleri

Önceden çekilen örnek birim mevcudu (n)	*k*= Çekilen ek birim sayısı													
	1	2	3	4	5	6	7	8	9	10	11	12	15	20
6	2.78	3.34	3.66	3.90	4.09	4.23	4.35	4.45	4.55	4.63	4.70	4.77	4.93	5.15
7	2.62	3.11	3.44	3.61	3.78	3.91	4.01	4.10	4.20	4.26	4.33	4.40	4.54	4.73
8	2.51	2.97	3.24	3.43	3.58	3.70	3.80	3.88	3.96	4.02	4.09	4.14	4.28	4.46
9	2.43	2.86	3.12	3.29	3.43	3.55	3.64	3.72	3.79	3.86	3.91	3.96	4.09	4.25
10	2.37	2.78	3.03	3.20	3.33	3.44	3.52	3.59	3.67	3.72	3.77	3.83	3.94	4.10
11	2.33	2.73	2.96	3.12	3.24	3.34	3.43	3.50	3.56	3.62	3.67	3.72	3.83	3.99
12	2.29	2.68	2.90	3.05	3.17	3.28	3.35	3.43	3.49	3.54	3.59	3.63	3.75	3.89
13	2.26	2.64	2.85	3.01	3.12	3.22	3.29	3.36	3.42	3.47	3.52	3.56	3.68	3.82
14	2.24	2.61	2.81	2.97	3.08	3.16	3.24	3.31	3.37	3.42	3.47	3.51	3.61	3.75
15	2.22	2.58	2.78	2.93	3.04	3.12	3.20	3.26	3.32	3.37	3.41	3.46	3.56	3.70
16	2.20	2.56	2.76	2.90	3.01	3.10	3.17	3.23	3.28	3.33	3.38	3.42	3.52	3.65
17	2.18	2.53	2.73	2.87	2.98	3.07	3.13	3.19	3.25	3.30	3.34	3.38	3.48	3.61
18	2.17	2.51	2.71	2.84	2.95	3.04	3.10	3.17	3.22	3.26	3.31	3.35	3.45	3.57
19	2.16	2.50	2.69	2.83	2.93	3.01	3.09	3.15	3.20	3.24	3.29	3.32	3.42	3.54
20	2.15	2.49	2.68	2.81	2.91	2.99	3.06	3.12	3.17	3.22	3.26	3.30	3.39	3.51
21	2.14	2.47	2.66	2.79	2.89	2.97	3.04	3.10	3.15	3.19	3.24	3.28	3.37	3.49
25	2.10	2.43	2.62	2.74	2.84	2.91	2.98	3.04	3.09	3.13	3.17	3.21	3.30	3.41
31	2.07	2.39	2.57	2.69	2.79	2.86	2.92	2.98	3.03	3.06	3.10	3.14	3.22	3.34
41	2.05	2.35	2.52	2.64	2.73	2.80	2.86	2.91	2.96	3.01	3.03	3.07	3.15	3.26
61	2.02	2.32	2.48	2.60	2.68	2.75	2.81	2.85	2.90	2.94	2.97	3.01	3.08	3.18
121	1.99	2.28	2.44	2.54	2.63	2.69	2.75	2.79	2.84	2.88	2.91	2.94	3.01	3.10
∞	1.96	2.24	2.39	2.50	2.58	2.64	2.69	2.74	2.78	2.82	2.84	2.88	2.95	3.04

Kaynak: Hahn Gerald J., "Factors For Calculating Two-Sided Prediction İntervals For Samples From A Normal Distribution, Jasa 64 (1969), 879.

Çizelge 4. Çift Taraflı %99 Öngörü Sınırları için k ve n Bileşimleri

Önceden çekilen örnek birim mevcudu (n)	*k*= Çekilen ek birim sayısı													
	1	2	3	4	5	6	7	8	9	10	11	12	15	20
6	4.36	5.08	5.51	5.82	6.07	6.27	6.44	6.58	6.71	6.83	6.93	7.02	7.26	7.57
7	3.96	4.56	4.93	5.19	5.39	5.56	5.69	5.82	5.93	6.02	6.10	6.18	6.38	6.64
8	3.71	4.24	4.56	4.78	4.96	5.10	5.23	5.34	5.43	5.51	5.58	5.65	5.83	6.05
9	3.54	4.02	4.30	4.51	4.66	4.80	4.90	5.00	5.08	5.16	5.23	5.29	5.45	5.65
10	3.41	3.85	4.11	4.31	4.45	4.57	4.67	4.76	4.83	4.90	4.96	5.02	5.16	5.35
11	3.31	3.73	3.97	4.14	4.28	4.39	4.49	4.57	4.64	4.71	4.76	4.82	4.95	5.12
12	3.23	3.63	3.86	4.03	4.15	4.26	4.35	4.43	4.49	4.55	4.60	4.65	4.78	4.95
13	3.17	3.55	3.77	3.93	4.05	4.15	4.23	4.31	4.37	4.43	4.48	4.52	4.65	4.80
14	3.12	3.48	3.70	3.86	3.96	4.06	4.15	4.21	4.27	4.33	4.38	4.42	4.53	4.69
15	3.07	3.43	3.64	3.78	3.90	3.98	4.07	4.13	4.19	4.24	4.29	4.34	4.44	4.58
16	3.04	3.38	3.59	3.73	3.84	3.92	4.00	4.07	4.13	4.18	4.21	4.26	4.36	4.50
17	3.01	3.34	3.54	3.68	3.78	3.87	3.94	4.00	4.06	4.11	4.16	4.20	4.30	4.43
18	2.98	3.31	3.50	3.63	3.73	3.82	3.89	3.95	4.01	4.06	4.10	4.14	4.24	4.37
19	2.95	3.28	3.46	3.59	3.70	3.78	3.85	3.91	3.96	4.01	4.06	4.09	4.19	4.32
20	2.93	3.26	3.44	3.56	3.66	3.74	3.81	3.87	3.92	3.97	4.01	4.05	4.14	4.27
21	2.91	3.23	3.41	3.54	3.63	3.71	3.77	3.83	3.88	3.93	3.98	4.02	4.10	4.23
25	2.85	3.15	3.32	3.44	3.53	3.61	3.67	3.73	3.78	3.81	3.86	3.89	3.98	4.09
31	2.79	3.08	3.24	3.35	3.44	3.51	3.57	3.62	3.66	3.71	3.74	3.78	3.86	3.97
41	2.74	3.01	3.16	3.27	3.35	3.42	3.48	3.52	3.56	3.60	3.64	3.66	3.74	3.84
61	2.68	2.94	3.08	3.19	3.26	3.33	3.38	3.42	3.46	3.50	3.52	3.56	3.63	3.72
121	2.63	2.88	3.01	3.11	3.17	3.24	3.28	3.33	3.36	3.40	3.43	3.45	35.2	3.61
∞	2.58	2.80	2.94	3.03	3.10	3.15	3.19	3.24	3.27	3.31	3.33	3.35	3.41	3.50

Kaynak: Hahn , 880.

Çizelge 2, 3, 4 ve (8) formülünün

$t[n-1; 1-(1-\gamma)/2k](n+1/n)^{1/2}$

kısmını sırasıyla $\gamma = 0.90$, $\gamma = 0.95$ ve $\gamma = 0.99$ için ikame etmektedir:

$u(n+1/n) = r(k,n;\gamma)$

Burada $r(k,n;\gamma)$ çeşitli γ değerlerine göre çizelgelerdeki k ve n bileşimlerini simgelemektedir. Böylece k için %100γ olasılıkla çift taraflı öngörü sınırları

$$\overline{Y} \pm r(k,n;\gamma)s \tag{19}$$

olmaktadır.

Tek taraflı öngörü sınırları için ise çizelge 5,6 ve 7'den benzer şekilde hareket edilmektedir.

Çizelge 5. Tek Taraflı %90 Öngörü Sınırları için k ve n Bileşimleri

Önceden çekilen örnek birim mevcudu (n)	1	2	3	4	5	6	8	10	12	15	20
					90% Öngörü sınırları						
4	1.83	2.48	2.87	3.15	3.36	3.54	3.81	4.02	4.18	4.39	4.64
5	1.68	2.24	2.57	2.80	2.98	3.12	3.34	3.52	3.66	3.82	4.04
6	1.59	2.11	2.40	2.64	2.76	2.89	3.09	3.25	3.37	3.52	3.71
7	1.54	2.02	2.29	2.48	2.63	2.75	2.93	3.08	3.19	3.33	3.50
8	1.50	1.96	2.22	2.40	2.54	2.65	2.83	2.96	3.07	3.20	3.36
9	1.47	1.92	2.17	2.34	2.47	2.58	2.75	2.87	2.98	3.10	3.26
10	1.45	1.88	2.13	2.29	2.42	2.53	2.69	2.81	2.91	3.02	3.18
11	1.43	1.86	2.09	2.26	2.38	2.48	2.64	2.76	2.85	2.97	3.11
12	1.42	1.84	2.07	2.23	2.35	2.45	2.60	2.71	2.81	2.92	3.06
15	1.39	1.79	2.01	2.26	2.28	2.37	2.52	2.62	2.71	2.82	2.95
20	1.36	1.75	1.96	2.11	2.21	2.30	2.44	2.54	2.62	2.72	2.85
25	1.34	1.72	1.93	2.07	2.18	2.26	2.39	2.49	2.57	2.67	2.79
30	1.33	1.71	1.91	2.05	2.15	2.24	2.36	2.46	2.54	2.63	2.75
40	1.32	1.96	1.89	2.02	2.12	2.20	2.33	2.42	2.50	2.59	2.70
60	1.31	1.67	1.86	1.99	2.09	2.17	2.29	2.38	2.46	2.54	2.65
∞	1.28	1.63	1.82	1.94	2.04	2.11	2.22	2.31	2.38	2.46	2.56

Kaynak: Hahn, "Additional...,",1671.

Çizelge 6. Tek Taraflı %95 Öngörü Sınırları için k ve n Bileşimleri

					95% Öngörü sınırları						
4	2.63	3.40	3.87	4.21	4.47	4.69	5.02	5.28	5.49	5.74	6.06
5	2.34	2.95	3.32	3.58	3.79	3.95	4.22	4.42	4.58	4.78	5.03
6	2.18	2.72	3.03	3.26	3.43	3.58	3.80	3.97	4.11	4.28	4.49
7	2.08	2.57	2.86	3.06	3.22	3.34	3.55	3.70	3.82	3.98	4.17
8	2.01	2.47	2.74	2.93	3.07	3.19	3.37	3.52	3.63	3.77	3.95
9	1.96	2.40	2.65	2.83	2.97	3.08	3.25	3.38	3.49	3.62	3.79
10	1.92	2.35	2.59	2.76	2.89	2.99	3.16	3.28	3.39	3.51	3.67
11	1.89	2.30	2.54	2.70	2.82	2.93	3.08	3.21	3.30	3.42	3.58
12	1.87	2.27	2.50	2.65	2.78	2.87	3.03	3.14	3.24	3.35	3.50
15	1.82	2.20	2.41	2.56	2.67	2.76	2.90	3.01	3.10	3.21	3.34
20	1.77	2.13	2.33	2.47	2.57	2.66	2.79	2.89	2.97	3.07	3.19
25	1.74	2.09	2.29	2.42	2.52	2.60	2.73	2.82	2.90	2.99	3.11
30	1.73	2.07	2.26	2.39	2.48	2.56	2.68	2.78	2.85	2.94	3.05
40	1.71	2.04	2.22	2.35	2.44	2.52	2.63	2.72	2.79	2.88	2.99
60	1.69	2.01	2.19	2.31	2.40	2.47	2.58	2.67	2.74	2.82	2.92
∞	1.64	1.95	2.12	2.23	2.32	2.39	2.49	2.57	2.63	2.71	2.80

Kaynak: Hahn "Additional...",1671.

Çizelge 7. Tek Taraflı %99 Öngörü Sınırları için k ve n Bileşimleri

99% Öngörü sınırları

4	5.07	6.30	7.07	7.63	8.07	8.43	9.00	9.43	9.79	10.22	10.76
5	4.10	4.94	5.46	5.83	6.13	6.37	6.75	7.04	7.28	7.57	7.95
6	3.63	4.30	4.70	4.99	5.22	5.41	5.71	5.94	6.12	6.35	6.64
7	3.36	3.93	4.27	4.51	4.70	4.86	5.11	5.30	5.46	5.65	5.90
8	3.18	3.69	3.99	4.20	4.37	4.51	4.73	4.90	5.04	5.21	5.43
9	3.05	3.52	3.79	3.99	4.14	4.27	4.46	4.62	4.74	4.90	5.09
10	2.96	3.39	3.65	3.83	3.97	4.09	4.27	4.41	4.53	4.67	4.85
11	2.89	3.30	3.54	3.71	3.84	3.95	4.12	4.25	4.36	4.50	4.67
12	2.83	3.22	3.45	3.61	3.74	3.84	4.00	4.13	4.23	4.36	4.52
15	2.71	3.07	3.27	3.42	3.53	3.62	3.77	3.88	3.97	4.08	4.22
20	2.60	2.93	3.11	3.24	3.34	3.42	3.55	3.65	3.73	3.83	3.95
25	2.54	2.85	3.02	3.14	3.24	3.31	3.43	3.52	3.60	3.69	3.80
30	2.50	2.80	2.97	3.08	3.17	3.24	3.36	3.44	3.51	3.60	3.71
40	2.46	2.74	2.90	3.04	3.09	3.16	3.27	3.35	3.41	3.49	3.59
60	2.41	2.68	2.83	2.94	3.02	3.08	3.18	3.26	3.32	3.39	3.49
∞	2.33	2.58	2.71	2.81	2.88	2.93	3.02	3.09	3.14	3.21	3.29

Kaynak: Hahn "Additional...",1671.

Çizelge 5,6 ve 7 bu kez $r'(k,n;\gamma) = u'(n+1)^{1/2}$ ifadesine karşılık gelmekte , alt ve üst öngörü sınırı

$$\bar{Y} - r'(k,n;\gamma)s \qquad (20)$$

$$\bar{Y} + r'(k,n;\gamma)s \qquad (21)$$

biçiminde gösterilmektedir.

Uygulama 1'deki problemin örnek birim sayısını 11'e çıkararak önce k=1, sonra da k=2 için çözüm aranırsa $(\gamma = 0.95)$:

Uygulama 2:

Y:10,9,7,10,3,9,12,5,8,13,6 $\quad \bar{Y} = 8.3636 \quad s = 2.9756 \quad (n+1/n)^{1/2} = 1.0445$

k=1 için (7) formülüne göre $8.3636 \pm (4.96)^{1/2}(1.0909091)^{1/2}(2.9756) = 8.3636 \pm 6.92$

$1.4436 < X < 15.28$

bulunmaktadır. (11) formülüne göre ise

$8.3636 \pm (2.228)(1.0909091)(2.9756) = 8.3636 \pm 6.92$

$1.4436 < X < 15.28$

olarak aynı sonuca erişilmektedir. Üçüncü olarak çizelge 1 kullanılarak öngörü sınırları hesaplanacaktır:

$8.3636 \pm (2.23)(1.0909091)^{1/2}(2.9756) = 8.3636 \pm 6.93$

$1.4336 < X < 15.294$

son olarak da (18) formülüne çizelge 3 verileri uygulanacaktır:

$8.3636 \pm (2.33)(2.9756) = 8.3636 \pm 6.93$

$1.4336 < X < 15.294$

Görüldüğü gibi dört yöntem ile aynı sonuca ulaşılmıştır. %95 olasılıkla çekilecek 12. birim X'in değeri 1.43 ila 15.29 arasında olacağı beklenmektedir.

Bu kez de k=2 için çözüm aranacaktır. $\gamma = 0.95$ ve k=2 için 2.64 verilmektedir. Buma göre de öngörü sınırları

$$8.3636 \pm (2.64)(1.0909091)^{1/2}(2.9756) = 8.3636 \pm 8.2049$$

$$0.1587 < X < 16.5685$$

bulunmaktadır. almaşık olarak (18) formülüne göre öngörü sınırları için

$$8.3636 \pm (2.73)(2.9756) = 8.3636 \pm 8.1234$$

$$0.2403 < X < 16.487$$

sonucuna ulaşılacaktır. Her iki yaklaşım karşılaştırıldığında özellikle alt hudutta bariz bir farklılık göze çarpmaktadır. Bunun nedeni ise çizelge 1'den hareket edildiğinde (2.64)(1.0444)=2.7572 kullanılırken çizelge 3'deki bileşimim kullanıldığından 2.73 sonucuna erişilmesidir.

Uygulama 3:

Otomobil satan bir şirket 121 hafta boyunca haftada ortalama 25 otomobil satmıştır. Dönemin standart sapması ise 5 olarak hesaplanmıştır. Yıl sonuna 2 hafta kaldığından şirket gelecek 2 hafta için %90 olasılıkla kaç otomobil satmayı bekleyebilir?
Bu problemde eldeki veya geçmişteki örneklerden hareketle birimlerin gelecekteki değerlerinin ne olacağı belirlenmek istenilmektedir.

Verilere; $\overline{Y} = 25$ n=121 k=2 s=5 $\gamma = 0,90$ $t[120, 0.975] = 2.27$ (8) formülü uygulanacak olursa:

$$\overline{Y} \pm t[n-1; 1-(1-\gamma)/2k](n+1/n)^{1/2} s$$

$$25 \pm 2.27(122/121)^{1/2} 5 = 25 \pm 11.40$$

$$13.6 < X < 36.4$$

%90 olasılıkla iki hafta boyunca, her hafta 14 ile 36 arasında otomobil satılması beklenilir.

Aynı problem için çizelge 2'den hareket edilirse bu kere (19)'den

$$\overline{Y} \pm (1.95).5 = 25 \pm 9.75$$

$$15.25 < X < 34.75$$

hesaplanacaktır. Gerek uygulama 1 gerekse uygulama 2 öngörü sınırlarının genişliği bakımından benzer sonuçlar vermektedir. Her ikisinde Hahn çizelgelerinin kullanıldığı (19) formülü daha dar sonuçlar vermiştir.

5. GENEL DEĞERLENDİRME VE SONUÇ

"Sınır" veya "Sınırlar" kavramları Türkçe yazında gerektiğinden çok daha dar bir kapsamda ele alınmaktadır. Tek değişkenli birimlerde μ ve σ^2 parametreleri için tek taraflı hipotez testleri yapılması halinde çift taraflı güven sınırları oluşturulmakta, tek taraflı güven sınırları ise hiç uygulanmamaktadır. Burada ise tek taraflı hipotez testinin karşıtının tek taraflı güven sınırı olduğu savunulmaktadır.

Diğer taraftan tek değişkenli veriler için "sınır" kavramının sadece güven sınırları ile ilintili olduğuna, öngörü sınırları, parametrik olmayan öngörü sınırlarına ve nihayet tolerans sınırlarına Türkçe yazında hiç yer verilmediğine değinilmiştir. Mühendislik bilim dallarına ilişkin olması, özellikle kalite kontrolünde uygulanması ve öngörü sınırları ile parametrik olmayan öngörü sınırlarından farklı [15] olması nedeniyle tolerans sınırları diğerlerinden ayrı tutulmaktadır.

Regresyon'da uzun zamandır kullanılmasına rağmen tek değişkenli verilerde kullanılmamasının bir eksiklik olduğu görüşünden hareketle yazıda tek değişkenli dağılımlarda değişkenler için de öngörü sınırlarının ana hatları ortaya konulmaya çalışılmıştır, çünkü daha önce çekilmiş bulunan bir örnekten elde edilen bilgiler çerçevesinde sonradan çekilebilecek örnek birimlerini yüksek bir olasılıkla kapsayacak bir aralık oluşturulması öngörü sınırlarını son derece önemli kılmaktadır.

Birimler arasındaki farkların fazla, dolayısıyla da değişkenliğin yüksek olduğu verilerde öngörü sınırları özellikle önem kazanmaktadır, çünkü bu gibi hallerde ortalama için güven sınırları ile sonradan çekilen birim için öngörü sınırları arasındaki fark artmakta bu da öngörü yerine güven sınırlarından hareket edildiğinde karar vermede yanıltıcı bir rol oynamaktadır.

Ancak bunca önemine rağmen ana kütlenin normal dağıldığı varsayımı ve öngörü sınırlarının normallikten sapmaya aşırı duyarlı olması konuya en belirgin eleştiridir. Nitekim ana kütle ortalaması için güven sınırlarında böyle bir durum söz konusu değildir.

Nihayet normallikten sapma halinde uygulanacak parametrik olmayan öngörü sınırlarına burada yer verilmemesi yazıya getirilecek belli başlı eleştirilerden biridir. Öneminden dolayı bu konunun başka bir yazıda inceleneceğini belirtmek acaba yeterli bir özür müdür?

KAYNAKLAR

[1] Mc Lean Alan, "The Predictive Approach To Teaching Statistics", Journal of Statistics Education, Vol 8, No 5, 2000.

[2] Vardeman Stephen B., "What About the Other Intervals", The American Statistician, August 1992, 193.

[3] Wonnacott R., Wonnacott T., "Introductory Statistics", John Wiley, New York, 1972, 210-211.
Siegel Andrew F., "Practical Business Statistics", Irwin Mc Graw Hill, Chicago, 1990, 294-296.

[4] Siegel Andrew F., "Practical Business Statistics", Irwin Mc Graw Hill, Chicago, 1990, 296.

[5] Işıkara Baki, "Regresyon Yöntemleri ve Sorunları", İstanbul Üniversitesi, İstanbul, Yayın No 2100, 1975,45.
Ertek Tümay, "Ekonometriye Giriş", O.D.T.Ü., Ankara, Yayın No: 22, 124, 1978,124.

[6] Hahn Gerald J., "Factors For Calculating Two-Sided Prediction Intervals for Samples From A Normal Distribution", Journal Of American Statistical Association, 878, September 1969, 878.

[7] Whitemore G.A., "Prediction Limits for a Univariate Normal Observation", The American Statistician, 141-143, May 1986,141-143.
Patel J.K., "Prediction Intervals A Review Communications in Statistics- Theory And Methods , 18, 2395-2465, 1989.2395-2365.
Scheuer E.M., "Let's Teach More About Prediction", Proceedings Of The Statistical Education Section, American Statistical Association, 133-137,1990.133-137.
Vardeman, 193-197.

[8] Siegel Andrew F., "Practical Business Statistics", Irwin Mc Graw Hill, Chicago, 1990, 297-299.
Steel Robert G. D., Torrie James H., Dickey David A.," Principles and Procedures of Statistics", Mc Graw Hill, New York, 62 –65.

[9] Scheffe H., "A Method For Judging All Contrasts in The Analysis of Variance", Biometrika, 40, 1953, 87-104.

[10] Hahn Gerald J., "Factors For Calculating Two-Sided Prediction Intervals for Samples From A Normal Distribution", Journal Of American Statistical Association, 878, September 1969, 879.

[11] Siegel Andrew F., "Practical Business Statistics", Irwin Mc Graw Hill, Chicago, 1990, 298.

[12] Siegel Andrew F., "Practical Business Statistics", Irwin Mc Graw Hill, Chicago, 1990, 339.

[13] Hahn Gerald J., "Factors For Calculating Two-Sided Prediction Intervals for Samples From A Normal Distribution", Journal Of American Statistical Association, 878, September 1969, 879-880.

[14] Hahn Gerald J., "Additional Factors For Calculating Prediction Intervals For Samples From A Normal Distribution". Journal Of American Statistical Association, 1671, December 1970.

[15] Preston Scott, "Teaching Prediction Intervals", Journal Of Statistics Education, Vol 8, No 3, 2000.