T.C.

ONDOKUZ MAYIS ÜNİVERSİTESİ

İSTATİSTİK BÖLÜMÜ

KESTİRİM YÖNTEMLERİ

Prof. Dr. Kamil ALAKUŞ

SAMSUN-2020

İÇİNDEKİLER

I. BÖLÜM: TAHMİN
II. BÖLÜM: TAHMİN METOTLARI VE NOKTA TAHMİN EDİCİLERİN ÖZELLİKLERİ
III. DOĞRUSAL REGRESYON VE EN KÜÇÜK KARELER TAHMİNİ
IV. DÜZELTME METOTLARI İLE TAHMİN

I. BÖLÜM: TAHMİN

I.1 Giriş

İstatistiklerin amacı, tek örneklem bilgisine dayanarak bir kitle hakkında sonuçlar çıkarmaktır. Zira kitleler parametre olarak isimlendirilen sayısal betimleme ölçüleriyle karakterize edilirler. Bu yüzden birçok istatistiksel araştırmanın amacı, bir kitleye ilişkin parametre(ler) hakkında değerlendirmeler yapmaktır. İstatistiksel değerlendirmelerin çoğu hem tahmin ve hem de hipotez testi işlemlerinden oluşur. Bu bölümün genel konusu, kitle parametrelerinin tahminidir. Örnekleme dağılımları buradaki tahmin işlemlerini geliştirmede önemli rol oynar.

Tahminin birçok pratik uygulamaları vardır. Örneğin; bir çamaşır makinesi üreticisi beş yıl garanti süresinden önce bozulan ilgili makinelerin bozulma oranı, π’yi tahmin etmek isteyebilir. Diğer önemli kitle parametreleri, kitle ortalaması, varyansı ve standart sapmadır. Örnek vermek gerekirse bir süper marketin gişeleri önünde ortalama bekleme süresi μ veya bir eletronik aletin ölçüm hatasının standart sapması, σ tahmin edilmek istenebilir. Bu ders süresinde ilgilenilen parametreler amaç parametresi olarak isimlendirilecektir.

Farz edilsin ki belli bir coğrafik bölgede bulunan maden filizinin bir ons (yaklaşık 31.1 gr) undan elde edilebilir ortalama cıva miktarı tahmin edilmek istensin. Bu iki farklı formda tahmin edilebilir. Örnek vermek gerekirse 0.15 ons gibi tek bir sayı olabilir. Burada amaç, bu sayının bilinmeyen kitle parametresine çok yakın olmasıdır. İşte bu tip tahminler **nokta tahmini** olarak isimlendirilir. Zira tek bir değer veya nokta μ’nün tahminidir. Diğer taraftan μ iki sayı arasında bulunur, örneğin; 0.12 ve 0.18 ons gibi. Bu son tahmin tipinde verilen bu iki değer $(0.12,\ 0.18)$ gibi bir aralık biçiminde gösterilebilir. Bu tip tahmin, **aralık tahmini** olarak isimlendirilir. Burada μ’nün mümkün değerlerinin bu aralığın içerisinde olduğu söylenir.

Bir nokta tahmin işlemi, amaç parametresini bir örneklemdeki bilgiyi bir tek değer veya noktaya ulaşmak için kullanır. Bir aralık tahmin işlemi ise örneklem bilgisini iki sayıya ulaşmak için kullanır ki ilgilenilen parametreyi içinde bulundurur. Her iki durumda da gerçek tahmin bir **tahmin edici** ile başarılır. Bu ise nokta veya aralık tahmin değer veya değerlerini bir örneklem verisinden nasıl tespit edileceğinin bir kuralıdır.

Tanım-I.1: Bir tahmin edici, bir örneklemdeki ölçüm değerlerine dayanarak bir tahminin nasıl hesaplanacağının kuralıdır.

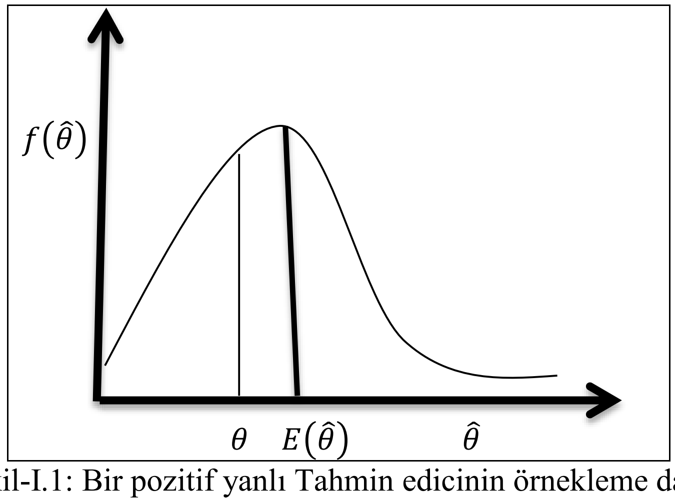
Çoğu kez bir tahmin edici, bir formülle ifade edilir. Örneğin; örneklem ortalaması, $\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$ kitle ortalaması μ’nün mümkün bir nokta tahmin edicisidir. Açık bir şekilde $\bar{Y}$ hem bir prensip ve hem de bir formüldür. Matematik olarak, gözlem değerlerinin toplamının örneklem hacmi, n'e bölünmesi anlamına gelir.

Bir parametreye ilişkin aralık tahmininde bulunmak isteyen bir araştırmacı iki nokta hesaplamak için örneklem verilerini kullanmalıdır. İki noktadan oluşan bu aralığın büyük bir olasılıkla amaç parametresini içinde bulunduracağı ümit edilir.

Aynı kitle parametresi için birçok farklı tahmin ediciler mümkündür. Bu sürpriz olmamalıdır. Örneğin; on mühendis büyük bir binanın yapım maliyetinin tahmininde her biri toplam maliyeti, farklı şekillerde tahmin edebilirler. İnşaat sanayiinde tahmin ediciler olarak isimlendirilen bu mühendisler bilinen yollar artı bitirme ön sezgilerini tahminlerinde kullanırlar. Her biri tek bir tahmin için tek bir insan sübjektifine dayalıdır. Bu düşünce, bazı tahmin edicilerin iyi, bazılarının da kötü olduğu fikri ile sonuçlanır. Öyle ise bir tahmin edicinin iyi mi kötü mü olduğu nasıl tespit edilecek? Bu düşüncenin cevabı hemen takipteki bölümde verilecektir.

I.2 Nokta Tahmin Edicilerin Bazı Özellikleri

θ ile gösterilen bir kitle parametresinin tahmin edicisi $\hat{\theta}$ ile gösterilsin. Tahminlerin dağılımı ve daha doğrusu tahmin edicinin örnekleme dağılımının bilinmesi gereklidir. Diğer bir ifade ile tahminlerin dağılımının ortalama veya beklenen değerinin bilinmesi gereklidir. Yani, $E(\hat{\theta}) = \theta$ yazılabilir. Bu özelliği sağlayan nokta tahmin edicilerin yansız olduğu söylenir. Pozitif yanlı (yani, $E(\hat{\theta}) > \theta$) bir nokta tahmin edicisinin örnekleme dağılımı, Şekil-I.1'de görüldüğü gibidir.



Şekil-I.1: Bir pozitif yanlı Tahmin edicinin örnekleme dağılımı.

Tanım-I.2: $\hat{\theta}$, θ'nın bir nokta tahmin edicisi olsun. Eğer , $E(\hat{\theta}) = \theta$ oluyorsa $\hat{\theta}$'ya θ için yansız (veya sapmasız) bir tahmin edicidir, denir. Aksi takdirde $\hat{\theta}$'nın yanlı olduğu söylenir.

Tanım-I.3: Bir $\hat{\theta}$ nokta tahmin edicisinin gerçek değerden yapması (yanlılığı), B ile gösterilir ve bu nicelik $B = E(\hat{\theta}) - \theta$ ile ifade edilir.

Yansızlığa ilave olarak tahminlerin örnekleme dağılımının yayılışının mümkün olduğu kadar küçük olması istenir. Yani, $V(\hat{\theta})$'nın en küçük olması istenir. Bir θ parametresi için iki yansız tahmin edici varsa ve diğer bütün şartlar aynı olmak şartıyla varyansı küçük olan tahmin edici diğerine yeğlenir.

Bir nokta tahmin edicisinin iyiliğini karakterize etmek için yan ve varyansı kullanmak yerine $\hat{\theta}$ ile θ arasındaki uzaklığın karesinin beklenen değeri kullanılır. Yani, $E\left[\left(\hat{\theta} - \theta\right)^2\right]$.

Tanım-I.4: Bir $\hat{\theta}$ nokta tahmin edicisinin hata kareler ortalaması, $MSE(\hat{\theta}) = E\left[\left(\hat{\theta} - \theta\right)^2\right]$ şeklinde tanımlanır, hem varyans ve hem de yanlılığın bir fonksiyonudur. Şöyle ki $MSE(\hat{\theta}) = V(\hat{\theta}) + B^2$ yazılabilir.

İspat: $\left(\hat{\theta} - \theta\right)^2$ ifadesine $E(\hat{\theta})$ ekleyip çıkarılırsa $\left(\hat{\theta} - E(\hat{\theta}) + E(\hat{\theta}) - \theta\right)^2 = \{[\hat{\theta} - E(\hat{\theta})] + E(\hat{\theta}) - \theta\}^2 = E\left[\left(\hat{\theta} - E(\hat{\theta})\right)^2\right] + E\left[\left(E(\hat{\theta}) - \theta\right)^2\right] = V(\hat{\theta}) + B^2$ olur.

UYGULAMA-I

Uygulama-I.1: Farz edilsin ki $E(\hat{\theta}_1) = E(\hat{\theta}_2) = \theta$ ve $V(\hat{\theta}_1) = \sigma_1^2$ ve $V(\hat{\theta}_2) = \sigma_2^2$ olsun. $\hat{\theta}_3$ yeni bir yansız tahmin edici olup $\hat{\theta}_3 = \alpha\hat{\theta}_1 + (1-\alpha)\hat{\theta}_2$ şeklinde tanımlansın.

a) $\hat{\theta}_1$ ve $\hat{\theta}_2$ tahmin edicilerinin bağımsızlığı varsayımı altında $\hat{\theta}_3$'ün varyansını en küçükleyecek α sabitini bulunur.

b) $\hat{\theta}_1$ ve $\hat{\theta}_2$ tahmin edicileri bağımsız değil ve $Cov(\hat{\theta}_1, \hat{\theta}_2) = c \neq 0$ ise $\hat{\theta}_3$'ün varyansını en küçükleyecek α sabitini bulunur.

Çözüm-I.1: a) $V(\hat{\theta}_3) = \alpha^2 V(\hat{\theta}_1) + (1-\alpha)^2 V(\hat{\theta}_2) = \alpha^2\sigma_1^2 + (1-\alpha)^2\sigma_2^2$ fonksiyonunu en küçük yapacak α sabitini bulmak için birinci türevi sıfırda değerlendirelim. $\frac{\partial V(\hat{\theta}_3)}{\partial \alpha} = 0 \Rightarrow 2\alpha(\sigma_1^2 + \sigma_2^2) - 2\sigma_2^2 = 0 \Rightarrow \alpha = \frac{\sigma_2^2}{(\sigma_1^2 + \sigma_2^2)}$ bulunur. b) $V(\hat{\theta}_3) = \alpha^2 V(\hat{\theta}_1) + (1-\alpha)^2 V(\hat{\theta}_2) = \alpha^2\sigma_1^2 + (1-\alpha)^2\sigma_2^2 + 2\alpha(1-\alpha)c$ fonksiyonunda bir kez türev alınırsa $\frac{\partial V(\hat{\theta}_3)}{\partial \alpha} = 0 \Rightarrow \alpha(\sigma_1^2 + \sigma_2^2 - 2c) - \sigma_2^2 + c = 0 \Rightarrow \alpha = \frac{\sigma_2^2 - c}{(\sigma_1^2 + \sigma_2^2 - 2c)}$ bulunur.

Uygulama-I.2: Y_1, Y_2 ve Y_3; $f(y) = \frac{1}{\theta}e^{-y/\theta}, y > 0$ olasılık yoğunluk fonksiyonuna sahip bir örneklemi göstersin. : $\hat{\theta}_1 = Y_1$, $\hat{\theta}_2 = \frac{Y_1 + Y_2}{2}$, : $\hat{\theta}_3 = \frac{Y_1 + 2Y_2}{3}$, $\hat{\theta}_4 = Enk\{Y_1, Y_2, Y_3\}$ ve $\hat{\theta}_5 = \bar{Y}$

θ'nın beş tahmin edicisi olmak üzere: a) Hangi tahmin edici(ler) yansızdır. b) a şıkkında belirlediğiniz yansız tahmin edici(ler)nin hangisi en küçük varyansa sahiptir.

Çözüm-I.2: a) $E(\hat{\theta}_1) = E(Y_1) = \theta$ olup yansızdır. $E(\hat{\theta}_2) = \frac{E(Y_1)+E(Y_2)}{2} = \frac{2\theta}{2} = \theta$ olup yansızdır. $E(\hat{\theta}_3) = \frac{E(Y_1)+2E(Y_2)}{3} = \frac{3\theta}{3} = \theta$ olup yansızdır. $E(\hat{\theta}_4) = \theta$ olup yansızdır. $E(\hat{\theta}_2) = \frac{E(Y_1)+E(Y_2)+E(Y_3)}{3} = \frac{3\theta}{3} = \theta$ yansızdır. Sonuç olarak, tahmin edicilerin hepsi de yansızdır. b) $V(\hat{\theta}_1) = V(Y_1) = \theta^2,\ V(\hat{\theta}_2) = \frac{V(Y_1)+V(Y_2)}{4} = \frac{2\theta^2}{4} = \frac{\theta^2}{2},\ V(\hat{\theta}_3) = \frac{V(Y_1)+4V(Y_2)}{9} = \frac{5\theta^2}{9},\ V(\hat{\theta}_4) = \theta^2$ ve $V(\hat{\theta}_5) = \frac{V(Y_1)+V(Y_2)+V(Y_3)}{9} = \frac{3\theta^2}{9} = \frac{\theta^2}{3}$ bulunur. O halde yansız tahmin edicilerin en küçük varyansa sahip olanı $\hat{\theta}_5 = \bar{Y}$'dır.

- ✓ **Not-I.1:** Ortalamadan fark kareler toplamı her zaman en küçüktür. (yani, $\sum_{i=1}^{n}(Y_i - \bar{Y})^2 = Enk$)

Uygulama-I.3: $f(y) = \frac{1}{\theta+1} e^{-\frac{y}{\theta+1}}, y > 0;\ \theta > -1$ foksinuna sahip bir kitlenin tesadüfi bir örneklemi; $Y_1, Y_2, \cdots, Y_n$ olsun. θ için yansız bir tahmin edici bulunuz.

Çözüm-I.3: $E(Y) = \int_0^{\infty} \frac{y}{(\theta+1)} e^{-y/(\theta+1)} dy = (\theta + 1)$ diğer taraftan en iyi tahmin edici ortalama olduğundan $\bar{Y} = \theta + 1$ eşitliğinden θ için yansız bir tahmin edici, $\hat{\theta} = \bar{Y} - 1$ olur.

Uygulama-I.4: Belli marka bilgisayarların heftalık arızalanlarının sayısı Y, λ parametreli bir Poisson tesadüfi değişken olsun. Böylesi bir kitlenin tesadüfi bir örneklemi; $Y_1, Y_2, \cdots, Y_n$ ise a) λ için yansız bir tahmin edici bulunuz. b) Arızalanmaların haftalık maliyeti, $C = 3Y + Y^2$ ise $E(C)$'nin $\lambda^2 + 4\lambda$ olduğunu gösteriniz. c) C için yansız bir tahmin edici tartışınız.

Çözüm-I.4: a) $P(Y = y) = p(y) = e^{-\lambda} \frac{\lambda^y}{y!}, y = 0,1,2,\cdots$ olup $E(Y) = \lambda => \hat{\lambda} = \bar{Y}$ yazılabilir. b) $E(C) = 3E(Y) + E(Y^2) = 3\lambda + \lambda^2 + \lambda = \lambda^2 + 4\lambda$ olur. (Not: Burda $Y \sim Pois(\lambda) =>\ E(Y) = V(Y) = \lambda$ olduğu unutulmamalıdır.) c) $E(C) = \lambda^2 + 4\lambda$ olduğundan $E(C)$'nın yansız tahmin edicisi, $\bar{Y}^2 + 3\bar{Y}$ olur.

Uygulama-I.5: Bir test devresine bağlı elektrik sayacının okumalarının $(\theta, \theta + 1)$ aralığında düzgün dağılımlı olduğu bilinmektedir. Burada θ devrenin bilinmeyen gerçek parametresidir. Bu tür okumaların tesadüfi bir örneklemi; $Y_1, Y_2, \cdots, Y_n$olsun. a) $\bar{Y}$, θ için yansız mıdır? Değilse yanı bulunuz? b) θ için yansız bir $\bar{Y}$ tahmin edicisi bulunuz. c) $\bar{Y}$, θ'nın bir tahmin edicisi olarak kullanıldığında $MSE(\bar{Y})$ bulunuz.

Çözüm-I.5: a) $(\theta, \theta + 1)$ aralığında düzgün dalılımlı bir Y tesadüfi değişkenin olasılıkk yoğunluk fonksiyonu, $f(y) = 1,\ \theta <\ y < (\theta + 1)$ olarak yazılır. $E(Y) = \int_{\theta}^{(\theta+1)} y dy =$

$\frac{y^2}{2}\Big|_{\theta}^{(\theta+1)} = \frac{[(\theta+1)^2-\theta^2]}{2} = \frac{(2\theta+1)}{2}$ olur. Öyle ise $E(\bar{Y}) = \frac{(2\theta+1)}{2} \neq \theta$ olup yanlıdır. $B = [E(\hat{\theta}) - \theta] = \frac{(2\theta+1)}{2} - \theta = \frac{1}{2} > 0$ olup pozitif yanlı bir tahmin edicidir. b) $\bar{Y} = \frac{(2\theta+1)}{2} => \bar{\theta} = \frac{(2\bar{Y}-1)}{2}$ tahmin edicisi θ için yansızdır. c) $MSE(\bar{Y}) = V(\bar{Y}) + B^2$ eşitliğinden hareketle, önce varyansı bulalım. $E(Y^2) = \int_{\theta}^{(\theta+1)} y^2 dy = \frac{y^3}{3}\Big|_{\theta}^{(\theta+1)} = \frac{[(\theta+1)^3-\theta^3]}{3} = \frac{(3\theta^2+3\theta+1)}{3}$ olup $V(Y) = E(Y^2) - [E(Y)]^2 = \frac{(3\theta^2+3\theta+1)}{3} - \frac{(4\theta^2+4\theta+1)}{4} = \frac{1}{12}$ bulunur. $V(\bar{Y}) = \frac{1}{12n}$'dir. Böylece $MSE(\bar{Y}) = V(\bar{Y}) + B^2 = \frac{1}{12n} + \left(\frac{1}{2}\right)^2 = \frac{(3n+1)}{12n}$ bulunur.

Uygulama-I.6: Olasılık yoğunluk fonksiyonu, $f(y) = \frac{\alpha y^{\alpha-1}}{\theta^{\alpha}}, 0 < y < \theta$ ile verilen bir kitleden *n* birimlik tesadüfi bir örneklem; $Y_1, Y_2, \cdots, Y_n$olsun. Bu fonksiyonda $\alpha > 0$ ve bilinen bir sabit; ancak θ bilinmemektedir. Eğer $\hat{\theta} = Enb\{Y_1, Y_2, \cdots, Y_n\}$ şeklinde verilmiş ise: a) $\hat{\theta}$, θ için yansız bir tahmin edici midir, gösteriniz? b) $MSE(\hat{\theta})$ bulunuz.

Çözüm-I.6: a) $E(Y) = \frac{\alpha}{\theta^{\alpha}} \int_0^{\theta} y^{\alpha} dy = \frac{\alpha y^{\alpha+1}}{(\alpha+1)\theta^{\alpha}}\Big|_0^{\theta} = \frac{\alpha\theta}{(\alpha+1)}$ bulunur. Buradan $E(\hat{\theta}) = E[Enb(Y_1, Y_2, \cdots, Y_n)] = \frac{\alpha\theta}{(\alpha+1)} \neq \theta$ olduğundan yanlı bir tahmin edicidir. b) $MSE(\hat{\theta}) = V(\hat{\theta}) + B^2$ olduğuna göre $B = E(\hat{\theta}) - \theta = \frac{\alpha\theta}{(\alpha+1)} - \theta = -\frac{\theta}{(\alpha+1)}$ bulunur. $=> B^2 = \frac{\theta^2}{(\alpha+1)^2}$ bulunur. $V(Y) = E(Y^2) - [E(Y)]^2$ olduğuna göre $E(Y^2) = \frac{\alpha}{\theta^{\alpha}} \int_0^{\theta} y^{\alpha+1} dy = \frac{\alpha y^{\alpha+2}}{(\alpha+2)\theta^{\alpha}}\Big|_0^{\theta} = \frac{\alpha\theta^2}{(\alpha+2)}$ olup $V(Y) = E(Y^2) - [E(Y)]^2 = \frac{\alpha\theta^2}{(\alpha+2)} - \left[\frac{\alpha\theta}{(\alpha+1)}\right]^2 = \frac{\alpha\theta^2}{(\alpha+1)^2(\alpha+2)}$ bulunur. Buradan $MSE(\hat{\theta}) = V(\hat{\theta}) + B^2 = \frac{2\theta^2}{(\alpha+1)(\alpha+2)}$ olarak elde edilir.

I.3 Bazı Önemli Yansız Nokta Tahmin Ediciler

Amaç parametrelerinin nokta tahmin edicilerini elde etmek için bazı biçimsel metotlar bir sonraki bölümde verilecektir. Bu bölümde ise bazı önemli nokta tahmin ediciler verilecektir. Örneğin; kitle ortalamasının (yani, μ'nün) tahmini için örneklem ortalamasının ($\bar{Y}$) kullanılması ve bir binomiyal parametresi π için örneklem oranı, $P = \frac{Y}{n}$'in kullanılması doğaldır. Öyle ise iki kitleden bağımsız olarak seçilmiş n_1 ve n_2 gözlemlerinin tesadüfi örneklemlerine dayanarak iki kitle ortalaması arasındaki farkı $(\mu_1 - \mu_2)$ veya iki binomiyal parametresinin farkı $(\pi_1 - \pi_2)$ nasıl tahmin edilecek? Yine $(\mu_1 - \mu_2)$ tahmini için iki örneklem ortalama farkı $(\bar{Y}_1 - \bar{Y}_2)$ ve $(\pi_1 - \pi_2)$ tahmini için de iki örneklem oran farkı $(P_1 - P_2)$ nokta tahmin edicileri olarak kullanılabilir.

Zira $\bar{Y}$, P; $(\bar{Y}_1 - \bar{Y}_2)$ ve $(P_1 - P_2)$ tahmin edicileri tesadüfi örnek ölçümlerinin fonksiyonlarıdır. Böylece onların beklenen değer ve varyansları beklenen değer teoremleri kullanılarak elde edilebilir. Tesadüfi örneklem(ler)e dayanarak bu dört tahmin ediciye ilişkin ortalama ve varyanslar Tablo-I.1'de verildiği gibidir.

Tablo-I.1'deki değerler şöyle elde edilebilir. Eğer örneklemler bağımsız ise $E(\bar{Y}) = \mu$ (tanımdan) ve $E(\bar{Y}_1 - \bar{Y}_2) = E(\bar{Y}_1) - E(\bar{Y}_2) = \mu_1 - \mu_2$ (tanımdan) gösterilebilir .$V(\bar{Y}_1 - \bar{Y}_2) = V(\bar{Y}_1) - V(\bar{Y}_2) - 2\underbrace{Cov(\bar{Y}_1, \bar{Y}_2)}_{(=0)} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$ (tanım ve bağımsızlıktan) gösterilebilir.

Tablo-I.1: Bazı nokta tahmin ediciler için beklenen değer ve varyanslar.

Amaç Parametresi (θ)	Örneklem Hacmi	Nokta Tahmin Edicisi $(\hat{\theta})$	$E(\hat{\theta})$	$V(\hat{\theta})$
μ	n	$\bar{Y}$	μ	$\frac{\sigma^2}{n}$
π	n	P	π	$\frac{\pi(1-\pi)}{n}$
$(\mu_1 - \mu_2)$	n_1 ve n_2	$(\bar{Y}_1 - \bar{Y}_2)$	$(\mu_1 - \mu_2)$	$\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$
$(\pi_1 - \pi_2)$	n_1 ve n_2	$(P_1 - P_2)$	$(\pi_1 - \pi_2)$	$\frac{\pi_1(1-\pi_1)}{n_1} + \frac{\pi_2(1-\pi_2)}{n_2}$

P ve $(P_1 - P_2)$ tahmin edicilerinin beklenen değer ve varyansı da ortalama veya ortalamalar farkına ilişkin yolla elde edilebilir.

Yansızlık, bir nokta tahmin edicisi için çoğu kez arzu edilen bir özellik olmasına rağmen her tahmin edici yansız değildir. Örneğin; $\tilde{S}^2 = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \bar{Y})^2$ tahmin edicisi kitle varyansı, σ^2 için yanlı bir tahmin edicidir. Genellikle bu yanlılık ortalamadan fark kareler toplamının n yerine $(n-1)$'e bölünmesiyle düzeltilebilir. Zira uygulamalarda S^2, σ^2'nin tahmini için çok sık kullanılmaktadır. Böylece σ^2'nin yansız nokta tahmin edicisi, $S^2 = \frac{1}{(n-1)}\sum_{i=1}^{n}(Y_i - \bar{Y})^2$ ile verilir ve bundan böyle ders boyunca örneklem varyansı olarak isimlendirilecektir.

Örnek-I.1: $Y_1, Y_2, \cdots, Y_n$ her biri μ ortalamalı ve σ^2varyanslı normal dağılımdan bir örneklem olsun. a) $\tilde{S}^2 = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \bar{Y})^2$ tahmin edicisinin σ^2 için yanlı, b) $S^2 = \frac{1}{(n-1)}\sum_{i=1}^{n}(Y_i - \bar{Y})^2$ tahmin edicisinin σ^2 için yansız olduğunu gösteriniz.

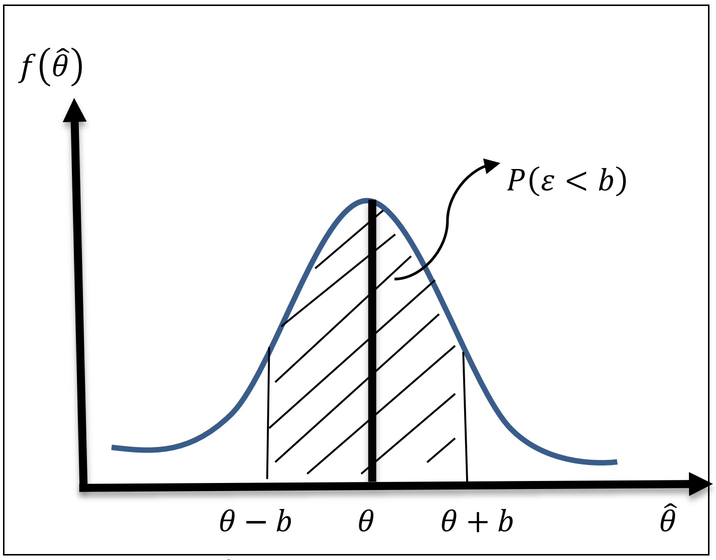
Çözüm-I.1: Önce payın beklenen değerini bulalım. Hemen $\sum_{i=1}^{n}(Y_i - \bar{Y})^2 = \sum_{i=1}^{n} Y_i^2 - n\bar{Y}^2$ yazılabilir. Buradan $E[\sum_{i=1}^{n}(Y_i - \bar{Y})^2] = E(\sum_{i=1}^{n} Y_i^2) - nE(\bar{Y}^2) = \sum_{i=1}^{n} E(Y_i^2) - nE(\bar{Y}^2)$ yazılabilir. Şimdi $E(Y_i^2)$, $i = 1,2,\cdots,n$ için $V(Y) = E(Y^2) - \mu^2$ ile verilen bir tesadüfi değişkenin varyansı durumuyla tamamen aynıdır. Öyle ise $E(Y^2) = V(Y) + \mu^2$ yazılabilir.

Böylece $E[\sum_{i=1}^{n}(Y_i-\bar{Y})^2]=E(\sum_{i=1}^{n}Y_i^2)-nE(\bar{Y}^2)=\sum_{i=1}^{n}(\sigma^2+\mu^2)-n\left(\frac{\sigma^2}{n}+\mu^2\right)=n(\sigma^2+\mu^2)-(\sigma^2+n\mu^2)=(n-1)\sigma^2$ sonucuna ulaşılır. Sonuç olarak; a) $E(\tilde{S}^2)=\frac{(n-1)}{n}\sigma^2\neq\sigma^2$ olduğundan yanlı, b) $E(S^2)=\frac{(n-1)}{(n-1)}\sigma^2=\sigma^2$ olduğundan yansızdır.

I.4 Bir Nokta Tahmin Edicinin İyiliğinin Değerlendirilmesi

Herhangi bir nokta tahmin edicisinin iyilik ölçüsününün bir yolu tahmin edici ile gerçek parametre arasındaki uzaklığı bulmaktır. Bu nicelik ise **tahmin hatası** olarak isimlendirilir. Doğal olarak tahmin hatasının mümkün olduğunca küçük olması arzu edilen bir sonuçtur.

Tanım-I.5: Tahmin hatası ε ile gösterilir ve tahmin edici ile onun bilinmeyen gerçek parametresi arasındaki uzaklıktır şeklinde tanımlanır. Yani $\varepsilon=|\hat{\theta}-\theta|$ ile verilir.



Şekil-I.2: Bir $\hat{\theta}$ Tahmin edicinin örnekleme dağılımı.

Zira tahmin hatası tesadüfi bir niceliktir ve belli bir tahmin için ne kadar büyük ve ne kadar küçük olacağı kesin söylenemez; ancak onun hakkında olasılık beyanında bulunulabilir. Örnek olarak Şekil-I.2'de gösterilen bir örnekleme dağılımına sahip $\hat{\theta}$, θ'nın yansız bir tahmin edicisi olsun. Eğer olasılık dağılımının uç noktalarına yakın bir yerde $(\theta-b)$ ve $(\theta+b)$ gibi iki nokta seçlirse tahmin hatasının olasılığı b'den daha küçüktür ve b ise Şekil-I.2'nin taralı bölgesidir. Yani $P(|\hat{\theta}-\theta|<b)=P\{-b<(\hat{\theta}-\theta)<b\}=P\{(\theta-b)<\hat{\theta}<(\theta+b)\}$ yazılabilir.

Burada b, tahmin hatasının bir olasılık sınırı gibi düşünülebilir. Buradan verilen bir hatanın b'den daha küçük olduğu kesin değildir; ancak böyle bir olayın büyük bir olasılığa sahip olduğu söylenebilir. Uygulama açısındaneğer b küçük seçilirse $P(\varepsilon<b)$ tek bir tahminin iyiliğinin bir ölçüsü gibi düşünülebilir.

Eğer $\hat{\theta}$'nın olasılık dağılımı bilinirse verilen bir tahmin problemi için b'nin değeri bulunabilir. %90 olasılıkla b'den daha küçük ε değeri bulunmak istendiğinde $\int_{\theta-b}^{\theta+b} f(\hat{\theta})\, d\hat{\theta} = 0.90$ eşitliğini sağlayan bir b değeri bulunabilir.

$\hat{\theta}$'nın olasılık dağılımı ister bilinsin ve isterse de bilinmesin yansız tahmin ediciler için yaklaşık bir hata sınırı; $b = k\sigma_{\hat{\theta}}$, $k \geq 1$ eşitliği kullanılarak elde edilebilir. Böylece Tchebysheff teoreminden bilinir ki ε'un $k\sigma_{\hat{\theta}}$'den daha küçük olması olasılığı, $1 - \frac{1}{k^2}$ olasılığına eşittir. Uygulamalarda uygun olan ve çek sık kullanılan k değeri ise 2'dir. Buradan %75 olasılıkla ε'un $b = 2\sigma_{\hat{\theta}}$'dan daha küçük olacağı söylenebilir. Bir çok tesadüfi değişken %95 civarinda bir olasılıkla $Ortalama \pm 2 \times Standart\ sapma$ arasında gözlemlenmektedir. Y'nin $(\mu \pm \sigma)$ aralığında bulunması olasılığı normal, düzgün ve üstel olasılık dağılımları için Tablo-I.2'de verilmiştir.

Tablo-I.2: Çeşitli Olasılık Dağılımları İçin $P\{(\mu - 2\sigma) < Y(\mu + 2\sigma)\}$ Değerleri.

Dağılım Adı	Olasılık
Normal	0.9544
Düzgün	1.0000
Üstel	0.9502

$b = 2\sigma_{\hat{\theta}}$ değeri uygulamalar için tahmin hatası hakkında iyi bir yaklaşımdır. Tahmin hatasının bu sınırdan daha küçük olması olasılığı yaklaşık olarak %95'dir.

Örnek-I.2: Bir şehirde belediye başkan adayı Mehmet Bey'e oy veren oranını tahmin etmek isteyen araştırmacı, tesadüfi olarak seçilen $n = 1000$ birimlik bir örneklemdeki seçmenlerin 575 tanesinin lehte oy kullanacakları tespit etmştir. a) Şehirdeki seçmenlerin Mehmet Bey'i seçme oranını tahmin ediniz. a) Tahmin hatasını $2\sigma_{\hat{\theta}}$ yapan b değerini bulunuz.

Çözüm-I.2: a) π'yi tahmin edebilmek için, $P = \frac{Y}{n}$ tahmin edicisi kullanılır. Buradan Mehmet Bey'e oy verenlerin oranı, $p = \frac{575}{1000} = 0.575$ olarak bulunur. b) P'nin olasılık dağılımı, $n = 1000$ olup yeteri kadar büyük hacimli örneklemler için normal dağılıma neredeyse kesin olarak yaklaşır. Öyle ise $b = 2\sigma_{\hat{\theta}}$ iken ε'un b'den daha küçük olması olasılığı yaklaşık olarak %95'dir. Tablo-I.1'den $V(P) = \frac{\pi(1-\pi)}{n}$'dir. Buradan $b = 2\sigma_{\hat{\theta}} = 2\sqrt{\frac{\pi(1-\pi)}{n}}$ yazılabilir.

Maalesef b'yi hesaplamak için π'nin bilinmesi gereklidir. π'nin bulunması örneklemenin amacı olmuştur. Bu aşılamaz bir engel değildir. Zira $\sigma_{\hat{\theta}}$, π'de çok az değişim gösterecektir. Buradan eşitlikteki π yerine P tahmininin (p) kullanılmasıyla $b = 2\sigma_{\hat{\theta}}$ tam değerin hesaplanmasında çok önemsiz bir hata olacaktır. Öyle ise bu örnek için $b = 2\sigma_{\hat{\theta}}=2\sqrt{\frac{PQ}{n}} = 2\sqrt{\frac{0.575(0.425)}{1000}} \cong 0.031$ olarak bulunur. Bu sonucun anlamı ise yaklaşık olarak %95 olasılıkla tahminin hatası 0.031'den daha küçüktür. Bu nedenle kitle oranı π'nin değerinin büyük bir güvenle 0.575 ± 0.031 sınırları içerisinde değişeceği söylenebilir.

Örnek-I.3: İki tip otomobil lastik tekerlerinin kalitesini karşılaştırmak amacı ile her bir lastik tipinden $n_1 = n_2 = 100$ adet lastik yol testine tabi tutulmuştur. Eskiyinceye kadar geçen mesafe kilometre olarak kaydedilmiştir. Test sonuçları: $\bar{y}_1 = 42500\ Km.$,$s_1^2 = 1440000Km^2$, $\bar{y}_2 = 40500\ Km.$ ve $s_2^2 = 1960000Km^2$ olarak hesaplanmıştır. a) Lastik tiplerinin eskime ortalamalar farkını tahmin ediniz. b) Tahmin hatasını içinde bulunduracak $b = 2\sigma_{\hat{\theta}}$ yapan değerini belirleyiniz.

Çözüm-I.3: a) $(\mu_1 - \mu_2)$'nin nokta tahmini, $(\bar{y}_1 - \bar{y}_2) = 42500 - 40500 = 2000$ Km olarak tahmin edilir. b) $\sigma_{\bar{Y}_1-\bar{Y}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$ değerini hesaplayabilmek için σ_1^2 ve σ_2^2 değerleri bilinmeli veya onların iyi yaklaşımlı değerleri bilinmelidir. σ_1^2 ve σ_2^2 değerleri daha önceki aynı konularda yapılan çalışmalardan da bilinebilir veya ilgili örneklem değerleri yardımıyla yansız tahmin ediciler kullanılarak elde edilebilirler. Şöyle ki $S_j^2 = \frac{1}{n_j-1}\sum_{i=1}^{n_j}\left(Y_{ji} - \bar{Y}_j\right)^2,\ j = 1,2$ tahmin ediciler yardımıyla hesaplanabilir. Eğer örneklem hacimleri yeterince büyükse (her biri için $n_j \geq 30$; $j = 1,2$) bu tahmin edicilerin ürettiği tahminler yeterlidir. Bu nicelikler verilen örnekte sırası ile 1440000Km^2 ve 1960000Km^2 olduğuna göre bu sonuçların σ_1^2 ve σ_2^2 yerine kullanılması ile $\sigma_{\bar{Y}_1-\bar{Y}_2} \cong \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = \sqrt{\frac{3400000}{100}} = 184$ Km olarak tahmin edilir. Sonuç olarak; yaklaşık %95 olasılıkla ortalamalar arası fark, $b = 2\sigma_{\bar{Y}_1-\bar{Y}_2} = 2(184) = 368$ Km'den daha küçüktür. Bu nedenle lastik tiplerinin kitle ortalama farklarına ilişkin $(\mu_1 - \mu_2)$'nin değerinin büyük bir güvenle 2000 ± 368 Km sınırları içerisinde değişeceği söylenebilir.

UYGULAMA-II

Uygulama-II.1: Bir jeolog bölgesel kaya yapısında bir çok çatlama bulunan bölgede yer kabuğundaki kaymalar neticesinde ayrılmayı araştırmak istemektedir. Çatlakların ortalama açısını elde etmek için $n = 50$ birimlik bir çatlak örneklemine ilişkin ortalama kaymanın 39.8^0 ve standart sapmanın ise 17.2^0 olduğunu tespit etmiştir. a) Yer çatlaklarının yönünün ortalama açısını tahmin ediniz. b) Tahmin hatasını içinde bulunduracak $b = 2\sigma_{\bar{Y}}$ değerini belirleyiniz.

Çözüm-II.1: a) $E(\bar{Y}) = \mu$ olup μ'nün tahmini, $\bar{y} = 39.8^0$'dir. b) $b = 2\sigma_{\bar{Y}}$ => $\sigma_{\bar{Y}} = \frac{s}{\sqrt{n}} = \frac{17.2}{\sqrt{50}} \cong 2.43^0$ olur. Buradan $b = 2(2.43) = 4.86^0$ bulunur. Böylece $(39.8 \pm 4.86)^0$ şeklinde yazılır.

Uygulama-II.2: Bir bölgeden tesadüfi olarak $n = 200$ birimlik bir örneklem tasarruf hesaplarından anlaşılmıştır ki geçmiş bir yıl üzerinde aylık ortalama olarak %7.2 artmıştır. Yine aynı örneklemin standart sapması %5.6 olarak hesaplanmıştır. a) Bölgedeki tasarruf artış ortalamasını tahmin ediniz. b) Tahmin için $b = 2\sigma_{\hat{\theta}}$ sınırı belirleyiniz.

Çözüm-II.2: a) $E(\bar{Y}) = \mu$ olup μ'nün tahmini, $\bar{y} = 0.072$'dir. b) $b = 2\sigma_{\bar{Y}}$ => $\sigma_{\bar{Y}} = \frac{s}{\sqrt{n}} = \frac{0.056}{\sqrt{200}} \cong 0.004$ olur. Buradan $b = 2(0.004) = 0.008$ bulunur. Böylece (0.072 ± 0.008) şeklinde yazılır.

Uygulama-II.3: Ortalaması θ olan bir üstel dağılımın bir tesadüfi örneklemi; $Y_1, Y_2, \cdots, Y_n$ olsun. Her bir $i = 1,2,\cdots,n$ için $E(Y) = \theta$ ve $V(Y) = \theta^2$ olduğuna göre $E(\bar{Y}) = \theta$ ve $V(\bar{Y}) = \frac{\theta^2}{n}$ yazılır. θ'nın bir yansız tahmin edicisini tartışınız ve tahmin edicinin standart sapmasını nasıl tahmin edeceğinizi gösteriniz.

Çözüm-II.3: $E(\bar{Y}) = \frac{1}{n}E(Y_1 + Y_2 + \cdots + Y_n) = \frac{1}{n}[E(Y_1) + E(Y_2) + \cdots + E(Y_n)] = \frac{1}{n}\underbrace{(\theta + \theta + \cdots + \cdots)}_{n-tane} = \frac{n\theta}{n} = \theta$ => $\hat{\theta} = \bar{Y}$ yazılabilir. $V(\bar{Y}) = \frac{\theta^2}{n}$ olup standart sapması, $\sqrt{V(\bar{Y})} = \frac{\theta}{\sqrt{n}} = \frac{\bar{Y}}{\sqrt{n}}$ tahmin edicisi ile tahmin edilir.

I.5 Güven Aralıkları

Hatırlananacağı gibi bir aralık tahmin edicisi aralığın uç noktalarından iki sayının örneklem ölçümlerini kullanan metodu açıklayan bir kuraldır. İdeal olarak iyi bir aralık tahmin edicisinde iki özellik aranır. Birincisi oluşturulacak aralığın bilinmeyen kitle parametresi θ'yı içermesi ve ikincisi ise aralığın mümkün oldukça dar olmasıdır.

Güven aralığının uç noktalarından biri veya ikisi örneklem ölçmlerinin bir fonksiyonu olduklarından örneklemden örnekleme değişkenlik gösterirler. Böylece boyları ve yeri tesadüfi nicekliktir ve dolayısıyla kesin değildir. Genellikle sabit bir kitle parametresi bir tek örneklemden hesaplanan herhengi bir aralığın uç noktaları arasında yer alır. Burada amaç, büyük bir olasılıkla kitle parametresini içine alacak dar aralıklar garanti eden aralık tahmin ediciler bulmaktır.

Aralık tahmin ediciler, genellikle **güven aralığı** olarak isimlendirilirler. Bir güven aralığının üst ve alt uç noktaları **üst ve alt güven limitleri** olarak isimlendirilir. Kitle parametresi θ'yı içine alan bir güven aralığının olasılığı ise **güvenilirlik düzeyi** adını alır. Tahmin edici ile birleştirilen güvenilirlik düzeyi büyük ise bu aralığın θ'yı içinde bulundurma olasılığının büyyük olduğu söylenir.

$\hat{\theta}_L$ ve $\hat{\theta}_U$, bir θ parametresi için sırasıyla alt ve üst güven limitleri olsun ve ayrıca $(1-\alpha)$ güvenilirlik düzeyi olmak üzere güven aralığı, $P\{\hat{\theta}_L < \theta < \hat{\theta}_U\} = (1-\alpha)$ ile verilir. $\hat{\theta}_L$'den $\hat{\theta}_U$'ya tanımlanan bu aralığa **çift yönlü** güven aralığı adı verilir.

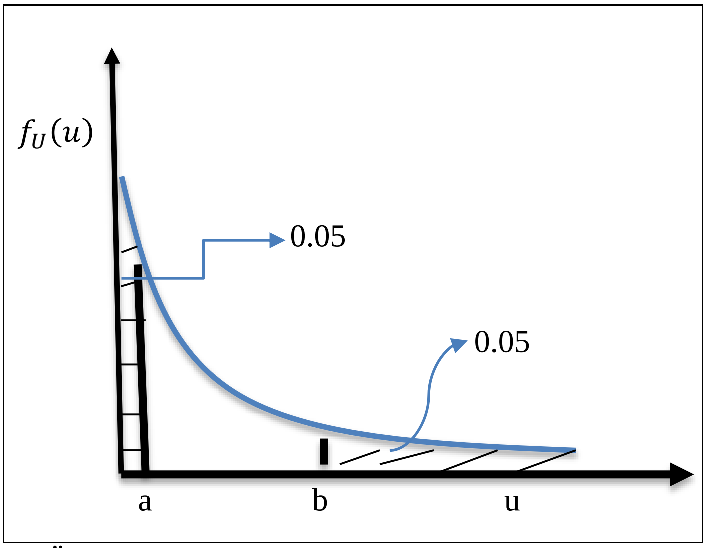
Ayrıca tek yönlü güven aralıkları da oluşturmak mümkündür. Belli bir güvenilirlikle alt sınır, $P\{\hat{\theta}_L < \theta\} = (1-\alpha)$ bulunur. Bu olayda sadece bir nokta tesadüfi olduğu halde güven aralığı, $(\hat{\theta}_L, \infty)$'dir. Benzer şekilde belli bir güvenilirlikle üst sınır, $P\{\theta < \hat{\theta}_U\} = (1-\alpha)$ bulunabilir. Bu anlamda güven aralığı ise $(-\infty, \hat{\theta}_U)$'dır.

Güven aralışlarının oluşturulmasında önemli bir metot **döngüsel metot** olarak isimlendirilir. Bu metot takipteki iki özelliği gerçekleyen döngüsel bir niceliğin bulunmasına bağlıdır. Bunlar: 1) Örnek ölçümlerinin ve bilinmeyen θ parametresinin bir fonksiyonudur. 2) Döngüsel niceliğin bir olasılık dağılımı varsa θ parametresine bağlı değildir.

Eğer döngüsel niceliğin olasılık dağılımı varsa (biliniyorsa) düşünülen aralık tahmini formu için takipteki mantık kullanılabilir. Y herhangi bir tesadüfi değişken, c bir sabit $(c > 0)$ ve $P(a \leq Y \leq b) = 0.75$ ise buradan $P(ca \leq cY \leq cb) = 0.75$ hemen yazılabilir. Benzer şekilde herhangi bir d sabiti için $P(a+d \leq Y+d \leq b+d) = 0.75$ yazılır. Yani bir olayın olasılığı, Y'nin bir dönüşüm ve ölçüsünün değişiminden etkilenmez. Böylece bir döngüsel niceliğin olasılık dağılımı bilinirse düşünülen aralık tahmin edici formlarında kullanıldığı gibi işlemleri kullanmak mümkün olabilir. Bu önemli metot takipteki iki örnek (Örnek-I.4 ve Örnek-I.5) üzerinde açıklanacaktır.

Örnek-I.4: Ortalaması θ olan bir üstel dağılımdan tek bir gözlem, Y düşünülsün. %90 güvenle θ için döngüsel metodu kullanarak bir güven aralığı bulunuz.

Çözüm-I.4: Y'nin olasılık yoğunluk fonksiyonu, $f(y) = \theta^{-1}exp(y/\theta), y > 0$ ile verilir. Döngüsel metodun kullanılmasıyla, $U = \frac{Y}{\theta}$ dönüşümü birim üstel dağılıma dönüşür. Böylece U'nun olasılık yoğunluk fonksiyonu, (DDT ile) a) Y'nin tanım bölgesi, $A = \{y: y > 0\}$, b) U'nun tanım bölgesi, $B = \{u: u > 0\}$, c) $J = \left|\frac{\partial y}{\partial u}\right| = \theta$ ve $f_U(u) = \exp(-u), u > 0$ olur ki birim (ya da standart) üstel dağılımın olasılık yoğunluk fonksiyonudur. U'nun olasılık yoğunluk fpnksiyonu Şekil-I.3'de grafik olarak gösterilmiştir.

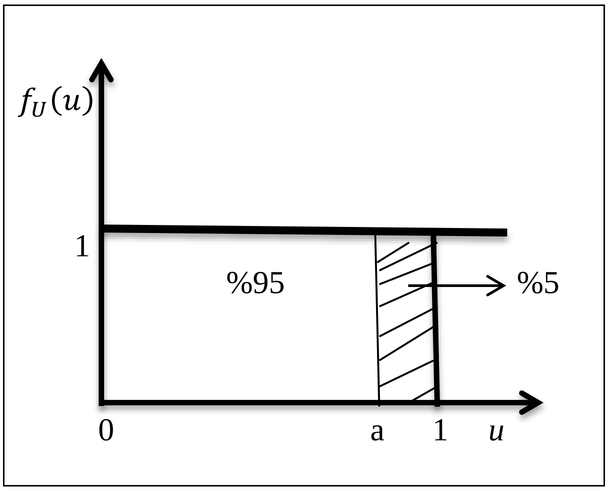


Şekil-I.3: Örnek-I.4'ün U dönüşümü için olasılık yoğunluk fonksiyonu.

$U = \frac{Y}{\theta}$, Y'nin ve θ parametresinin bir fonksiyonudur. Buradan U'nun dağılımı θ'dan bağımsızdır. Böylece $U = \frac{Y}{\theta}$ niceliği bir döngüsel nicelik olarak kullanılabilir. Zira %90 güvenilirlikle bir aralık tahmin edicisi a ve b $P\{a \leq U \leq b\} = 0.90$ ile verilebilir. a ve b'nin belirlenmesinde kullanılan yollardan biri $P(u \leq a) = \int_0^a e^{-u} du = 0.05$ ve $P(u \geq b) = \int_b^\infty e^{-u} du = 0.05$ eşitliklerinin sağlanmasıdır. Bu eşitlikler $1 - e^{-a} = 0.05$ ve $e^{-b} = 0.05$ ile verilir ve bu eşitliklerin çözümü ise $a = -\ln(0.95) \cong 0.051$ ve $b = -\ln(0.05) \cong 2.996$dır. Buradan %90 güven aralığı ise $P\{0.051 \leq U \leq 2.996\} = P\left\{0.051 \leq \frac{Y}{\theta} \leq 2.996\right\} = 0.90$ ile verilebilir. Burada amaç, θ için güven aralığı oluşturmak olduğundan son ifadeye gerekli dönüşümler uygulandığında aynı eşitlik, $P\left\{\frac{Y}{2.996} \leq \theta \leq \frac{Y}{0.051}\right\} = 0.90$ şekline dönüşür. Böylece $\frac{Y}{2.996}$ ve $\frac{Y}{0.051}$ aranan alt ve üst güven limitleridir. Bu limitlerin sayısal değerlerini elde etmek için Y'nin bir gerçek değerinin elde edilmesi gereklidir. Bu değer aralık tahmin edilmesinde yerine yazılarak bir sayısal aralık elde edilebilir. Yani %90 güvenilirlikle θ'nın bilinmeyen gerçek değeri $(\frac{Y}{2.996}, \frac{Y}{0.051})$ aralığındadır.

Örnek-I.5: $(0, \theta)$ aralığında tanımlı bir sürekli düzgün dağılımdan tek bir örnek Y seçilsin. Burada θ bilinmemektedir. θ için %95 güvenilirlikle bir alt limit bulunuz.

Çözüm-I.5: Y, $(0, \theta)$ aralığında düzgün dağılımlı olduğuna göre değişken değiştirme tekniği kullanılmalıdır. Öyle ise $U = \frac{Y}{\theta}$ değişkeni $(0, 1)$'de düzgün dağılıma sahiptir. Yani, $f_U(u) = 1, 0 < u < 1$ olur. Şekil-I.4'de U'nun olasılık yoğunluk fonksiyonu sergilenmektedir.



Şekil-I.4: Örnek-I.5’in U dönüşümü için olasılık yoğunluk fonksiyonu.

Bu örnekte de U dönüşümü bir döngüsel niceliğin gereklerini yerine getirmektedir. Öyle ise θ için bir %95 alt güven limiti oluşturulacağına göre a’nın değeri, $P(a > U) = 0.05$ veya $P(U \leq a) = 0.95$ yazılabilir. Yani; $\int_0^a du = 0.95$ olup $a = 0.95$ bulunur. Böylece $P(U \leq 0.95) = P\left(\frac{Y}{\theta} \leq 0.95\right) = P(Y \leq 0.95\theta) = P\left(\frac{Y}{0.95} \leq \theta\right) = 0.95$ olarak bulunmuş olur. $\frac{Y}{0.95}$, %95 güvenilirlikle θ için bir alt güven limitidir. Y’nin herhangi bir gözlem değeri θ’dan daha küçük olacağından θ için bulunacak alt sınır yeterince uygundur ve bu değer Y’nin gözlem değerinden $(\frac{1}{0.95})$ kat daha büyüktür.

- ✓ **Not-2:** Gerçek değişken için alt limit, dönüşüm değişkeni için üst limit ve gerçek değişken için üst limit, dönüşüm değişkeni için ise alt limit bulunarak elde edilir.

UYGULAMA-III

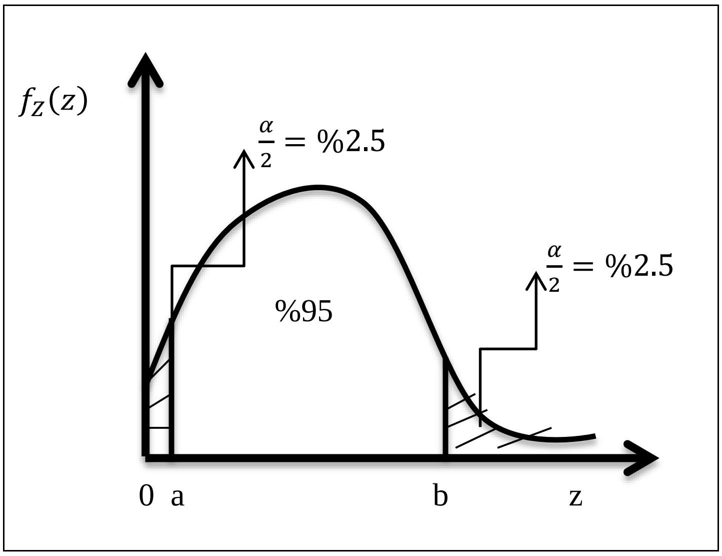
Uygulama-III.1: Y, ortalaması μ (bilinmeyen) ve varyansı 1 olan bir normal dağılımın bir örneği olsun. a) μ için bir %95 güven aralığı bulunuz. b) μ için bir %95 üst güven limiti bulunuz. c)) μ için bir %95 alt güven limiti bulunuz.

Çözüm-III.1: a) $Z_{\frac{\alpha}{2}} = \frac{(Y-\mu)}{(\sigma=1)} => \alpha = 0.05$ ve $\frac{\alpha}{2} = 0.025 => Z_{\frac{\alpha}{2}} = 1.96$ (z tablosundan) $=>$ $Y - \mu = 1.96$ yazılır. Böylece $\mu = (Y \pm 1.96)$ veya $P\{(Y - 1.96) \leq \mu \leq (Y + 1.96)\} = 0.95$ olarak elde edilir. b) $Z_\alpha = \frac{(Y-\mu)}{(\sigma=1)} => \alpha = 0.05$ için $z_\alpha = 1.645$ olup alt limit, $P\{(Y - 1.645) \leq \mu\} = 0.95$ olur. c) $Z_\alpha = \frac{(Y-\mu)}{(\sigma=1)} => \alpha = 0.05$ için $z_\alpha = -1.645 = Y - \mu$ olup üst limit, $P\{\mu \leq (Y + 1.645)\} = 0.95$ yazılır.

Uygulama-III.2: Y, ortalaması 0 ve bilinmeyen bir varyans σ^2 olan normal dağılımın bir örneği olsun. Öyle ise $Z = \frac{Y^2}{\sigma^2}$ dönünüşüm değişkeni 1 serbestlik dereceli Ki-Kare dağılımına sahiptir.

Bu dönüşümü kullanarak: a) σ^2 için bir %95 güven aralığı bulunuz. b)) σ^2 için bir %95 üst güven limiti bulunuz. c) σ^2 için bir %95 alt güven limiti bulunuz.

Çözüm-III.2: Bir serbestlik dereceli ki-kare dağılımlı Z dönüşüm değişkeninin olasılık yoğunluk fonksiyonu takipteki şekil ile verilir.



$P(a \leq Z \leq b) = 0.95$ yazılır. Buradan $P(Z \leq a) = \int_0^a f_Z(z)dz = 0.025 \Rightarrow \chi^2_{1;0.025} = 0.000982$ (ki-kare tablosundan) ve $P(Z \leq b) = \int_0^a f_Z(z)dz = 0.975 \Rightarrow \chi^2_{1;0.975} = 5.023886$ (ki-kare tablosundan) bulunur. $P(0.000982 \leq Z \leq 5.023886) = 0.95 \Rightarrow P\left(0.0009821 \leq \frac{Y^2}{\sigma^2} \leq 5.023886\right) = 0.95 \Rightarrow P\left(\frac{Y^2}{5.023886} \leq \sigma^2 \leq \frac{Y^2}{0.0009821}\right) = 0.95$ olarak elde edilir. b) σ^2 için bir %95 üst güven limit, Z dönüşümü için bir alt limit olduğundan $\chi^2_{1;0.05} = 0.003932$ (ki-kare tablosundan) bulunur. Gerekli düzenlemeler yapılırsa aranan güven aralığı, $P\left(\sigma^2 \leq \frac{Y^2}{0.003932}\right) = 0.95$ bulunmuş olur. c) σ^2 için bir %95 alt güven limit, Z dönüşümü için bir üst limit olduğundan $\chi^2_{1;0.95} = 3.841459$ (ki-kare tablosundan) bulunur. Gerekli düzenlemeler yapılırsa aranan güven aralığı, $P\left(\frac{Y^2}{3.841459} \leq \sigma^2\right) = 0.95$ bulunur.

Uygulama-III.3: Uygulama-III.2'nin veri ve sonuçlarını kullanarak σ için aynı istenenleri tekrarlayınız.

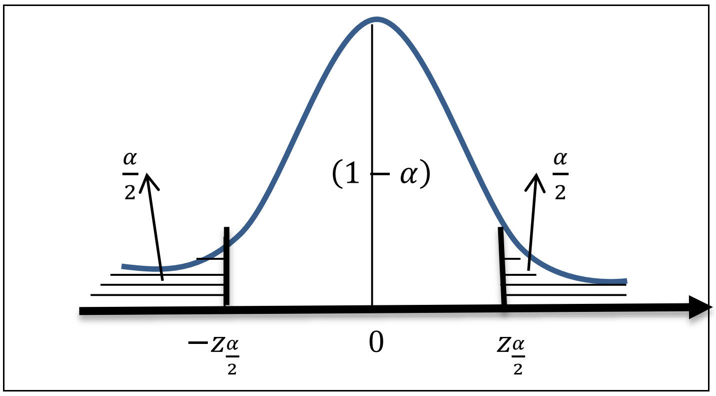
Çözüm-III.3: a) $P\left(\frac{Y}{\sqrt{5.023886}} \leq \sigma \leq \frac{Y}{\sqrt{0.0009821}}\right) = 0.95$ olur. b) $P\left(\sigma \leq \frac{Y}{\sqrt{0.003932}}\right) = 0.95$ ve c) $P\left(\frac{Y}{\sqrt{3.841459}} \leq \sigma\right) = 0.95$ olarak kolayca yazılabilir.

I.6 Büyük Hacimli Örneklemler İçin Güven Aralığı

Daha önceki bölümlerde μ, π, $(\mu_1 - \mu_2)$ ve $(\pi_1 - \pi_2)$ parametreleri için yansız tahmin ediciler verilmişti. Büyük hacimli örnek veya örneklemler için bu tahmin edicilerin hepsinin örnekleme dağılımı normaldir. Yani verilen şartlar altında bilinmeyen kitle parametresi θ, μ, π, $(\mu_1 - \mu_2)$ ve $(\pi_1 - \pi_2)$ ise buradan büyük hacimli örneklemler için $Z = (\hat{\theta} - \theta)\sigma_{\hat{\theta}}^{-1}$ yaklaşık olarak bir standart normal dağılımlı olur. Buradan $Z = \sigma_{\hat{\theta}}^{-1}(\hat{\theta} - \theta)$ bir döngüsel nicelik formundadır ve böylece döngüsel metot bilinmeyen kitle parametresi θ için aralık tahmin edicileri geliştirmede kullanılabilir.

Örnek-I.6: $\hat{\theta}$, ortalaması θ ve varyansı $\sigma_{\hat{\theta}}^2$ olan normal dağılımdan bir tahmin edici olsun. $(1 - \alpha)$ güvenilirlikle θ için güven aralığı bulunuz.

Çözüm-I.6: $Z = \sigma_{\hat{\theta}}^{-1}(\hat{\theta} - \theta)$ niceliği standart normal dağılıma sahip olduğuna göre bu dağılımın iki uç değeri, $-z_{\frac{\alpha}{2}}$ ve $z_{\frac{\alpha}{2}}$ seçilir. Buradan $P\left(-z_{\frac{\alpha}{2}} \leq Z \leq z_{\frac{\alpha}{2}}\right) = (1 - \alpha)$ yazılabilir. Bu durum Şekil-I.5'de gösterildiği gibi açıklanabilir.



Şekil-I.5: $-z_{\frac{\alpha}{2}}$ ve $z_{\frac{\alpha}{2}}$ yeri.

Son eşitlikteki Z'nin yerine ifadesinin yazılmasıyla, $P\left(-z_{\frac{\alpha}{2}} \leq \sigma_{\hat{\theta}}^{-1}(\hat{\theta} - \theta) \leq z_{\frac{\alpha}{2}}\right) = (1 - \alpha)$ elde edilir. Eşitliğin son tarafındaki her üç ifadenin $\sigma_{\hat{\theta}}$ ile çarpımı sonucu $P\left(-z_{\frac{\alpha}{2}}\sigma_{\hat{\theta}} \leq (\hat{\theta} - \theta) \leq z_{\frac{\alpha}{2}}\sigma_{\hat{\theta}}\right) = (1 - \alpha)$ elde edilir. Bu son ifadenin sol üç teriminden $\hat{\theta}$'nın çıkarılmasıyla $P\left(-z_{\frac{\alpha}{2}}\sigma_{\hat{\theta}} - \hat{\theta} \leq -\theta \leq z_{\frac{\alpha}{2}}\sigma_{\hat{\theta}} - \hat{\theta}\right) = (1 - \alpha)$ olur. Nihayet sol üç terimin (-1) ile çarpılması ve eşitsizliğin yönünün değiştirilmesiyle aranan güven aralığı, $P\left(\hat{\theta} - z_{\frac{\alpha}{2}}\sigma_{\hat{\theta}} \leq \theta \leq \hat{\theta} + z_{\frac{\alpha}{2}}\sigma_{\hat{\theta}}\right) = (1 - \alpha)$ elde edilir. Böylece θ için alt ve üst güven limitleri sırasıyla: $LCL = \hat{\theta} - z_{\frac{\alpha}{2}}\sigma_{\hat{\theta}}$ ve $UCL = \hat{\theta} + z_{\frac{\alpha}{2}}\sigma_{\hat{\theta}}$ ile verilir.

Bu örnek üzerinde anlatılan güven aralığı işlemi büyük hacimli örneklemler için μ, π, $(\mu_1 - \mu_2)$ ve $(\pi_1 - \pi_2)$ parametrelerinin güven aralıklarının bulunmasında kullanılabilir.

Örnek-I.7: Bir bölgenin süpermarketi müşterileri arasından tesadüfi olarak seçilen 64 tanesinin alış-veriş süreleri kaydedilmiştir. Bu 64 müşteriye ilişkin ortalama alış-veriş süresi ve varyansı sırasıyla 33 dk ve 256 dk^2 olarak hesaplanmıştır. %90 güvenilirlikle gerçek alış-veriş süresi μ için bir güven aralığı tahmin ediniz.

Çözüm-I.7: $\hat{\theta} = \bar{y} = 33\ dk$, $s^2 = 256$ dk^2 ve $n = 64$ müşteri olduğuna göre kitle varyansı σ^2 bilinmiyor; ancak $n \geq 30$ olduğu için tahmin değeri s^2 kullanılabilir. Aranan güven aralığı, $\hat{\theta} \pm z_{\frac{\alpha}{2}}\sigma_{\hat{\theta}}$ veya $\bar{y} \pm z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \approx \bar{y} \pm z_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}}$ formuna sahip olacaktır. Standart normal dağılım tablosundan $z_{\frac{\alpha}{2}} = z_{0.05} = 1.645$ bulunur. Buradan güven sınırları (limitleri), alt sınır için $LCL = \bar{y} - z_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}} = 33 - 1.645\left(\frac{16}{8}\right) = 29.71$dk. ve üst sınır için ise $LCL = \bar{y} + z_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}} = 33 + 1.645\left(\frac{16}{8}\right) = 36.29$ dk. olarak bulunur. Böylece μ'nün %90 güvenilirlikle güven aralığı, $(29.71, 36.29)$ yazılabilir.

Örnek-I.8: A ve B ile gösterilen iki marka buzdolabı beş yıl garanti belgelidir. A marka buzdolabından 50 tesadüfi örnek içinden 12 tanesi beş yıllık garanti süresi tamamlanmadan arızalanmıştır. B marka buzdolabından 60 tesadüfi örnek içinden ise yine 12 tanesi beş yıllık garanti periyodu içerisinde arızalanmıştır. %98 güvenle garanti periyodu içerisinde arıza oranları arasındaki gerçek fark, $(\pi_A - \pi_B)$ için güven aralığı tahmin ediniz.

Çözüm-I.8: Verilen değerler $\hat{\theta} \pm z_{\frac{\alpha}{2}}\sigma_{\hat{\theta}}$ ifadesinde değerlendirilirse aranan güven aralığı tahmin edicisi, $(P_A - P_B) \pm z_{\frac{\alpha}{2}}\sqrt{\frac{[\pi_A(1-\pi_A)]}{n_A} + \frac{[\pi_B(1-\pi_B)]}{n_B}}$ formuna sahip olur. π_A ve π_B değerleri bilinmediği için $\sigma_{\hat{\theta}}$'nın gerçek değeri elde edilemez; ancak büyük örnekli olma özelliğinden yararlanarak π_A ve π_B'nin yerine P_A ve P_B tahmin edicilerinin kullanılmasıyla $\sigma_{\hat{\theta}}$ için iyi bir tahmin elde edilebilir. Bu örnek için $p_A = \frac{12}{50} = 0.24$ ve $p_B = \frac{12}{60} = 0.20$ ve $z_{\frac{\alpha}{2}} = z_{0.01} = 2.33$ bulunur. Öyle ise aranan güven aralığı, $(0.24 - 0.20) \pm 2.33\sqrt{\frac{0.24(076)}{50} + \frac{0.20(0.8)}{60}}$ ya da 0.04 ± 0.1852 bulunur ve $(-0.15,\ 0.23)$ olarak yazılır.

- ✓ **Not-I.3:** Kitle oranları arasındaki farkın gerçek parametresi pozitif ya da negatif olduğundan aralığın alt sınırının negatif olması sürpriz değildir. Bu oranlar arasında fark olmadığı anlamına da gelir (zira aralık sıfırı içermektedir).

UGULAMA-IV

Uygulama-IV.1: Bir hastanenin yöneticileri 25 ila 34 yaşları arasındaki hastaların tedavileri için gerekli olan ortalama gün sayısını tahmin etmek istemektedir. Sözü edilen özelliğe sahip hastalar arasından 500 birimlik bir örnekleme ilişkin ortalama tedavi süresi, 5.4 gün ve standart sapma, 3.1 gün olarak hesaplanmıştır. Bilinmeyen kitle ortalaması için %95 güven aralığı tahmin ediniz.

Çözüm-IV.1: Verilen değerler $\hat{\theta} \pm z_{\frac{\alpha}{2}}\sigma_{\hat{\theta}}$ ifadesinde değerlendirilirse aranan güven aralığı, $\bar{y} \pm z_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}}$ formuna dönüşür. $\bar{y} = 5.4,\ s = 3.1\ n = 500$ ve $z_{0.025} = 1.96$ olduğuna göre %95 güvenle istenen aralık: $5.4 \pm 1.96(0.1386) = (5.12,\ 5.67)$ olarak bulunur.

Uygulama-IV.2: "A Üniversitesinin hukuk fakültesinde okuyan öğrencileri idam cezasına karşıdır" rapor sonucundan yola çıkarak yapılan bir araştırmada ilgili üniversitenin hukuk fakültesinde okuyan öğrenciler arasından tesadüfi olarak seçilen 86 öğrenci ile görüşme yapılmıştır. Görüşme sonucunda öğrencilerin %52'sinin idam cezasına karşı oldukları ortaya çıkmıştır. Bu bilgileri kullanarak ilgili üniversitenin hukuk fakültesi öğrencilerinin gerçek idam cezası karşı olanların oranı için %95 güven aralığı tahmin ediniz.

Çözüm-IV.2: Büyük hacimli örneklemler için kullanılan $\hat{\theta} \pm z_{\frac{\alpha}{2}}\sigma_{\hat{\theta}}$ ifadesi yardımıyla $\hat{\theta} = P$ olup tahmin değeri ise $p = 0.52$; $z_{0.025} = 1.96$ ve $s_P = \sqrt{\frac{0.52(0.48)}{86}} \cong 0.0539$ değerleri ilgili ifadede değerlendirilirse $0.52 \pm 1.96(0.0539) = (0.41,\ 0.63)$ aralığı bulunur.

Uygulama-IV.3: Doktorlar, he gün 50 ila 200 miligram selenyum elementinden alınmasının insan sağlığı açısından gerekli olduğunu ileri sürmektedir. Bu amaçla iki farklı bölgeden tesadüfi olarak 30'ar birimlik bireyle çalışma yapılmıştır. Bir günlük hem katı ve hemde sıvı halde selenyum verilmiş olup her bir birey için bu değerler kaydedilmiştir. Birinci bölgeden seçilen 30 erişkin bireye ait günlük alınan selenyum ortalama ve standart sapma sırasıyla 167.1 mg ve 24.3 mg olarak elde edilmiştir. İkinci bölgeden seçilen 30 erişkin bireye ilişkin aynı değerler sırasıyla 140.9 mg ve 17.6 mg olduğuna göre iki bölgenin alınan selenyum ortalama farkı için %95 güven aralığı tahmin ediniz.

Çözüm-IV.3: Verilen değerler $(\bar{Y}_1 - \bar{Y}_2) \pm z_{\frac{\alpha}{2}}\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$ ifadesinde değerlendirilirse aranan farka ilişkin güven aralığı, $(167.1 - 140.9) \pm 1.96\sqrt{\frac{(24.3)^2}{30} + \frac{(17.6)^2}{30}} = 26.2 \pm 1.96(5.48) = (15.46,\ 36.94)$ olarak elde edilir.

I.7 Örneklem Hacminin Belirlenmesi

Deney düzeni aslında bir bilgi niceliğinin elde edilmesi için bir plandır. Bu bilgi temin edilen verilerin durumuna bağlı olarak değişik biçimlerde elde edilebilir. Bazı ölçümler ilgilenilen parametreye ilişkin bilginin büyük miktarını içerirler; oysa diğerleri belki biraz belki de hiç gerektirmezler. Zira araştırmanın tek ürünü bilgidir ve böylece bu bilgiyi elde etme maliyeti en küçük ypılmalıdır.

Çoğu kez deney düzeni olarak isimlendirilen örnekleme yöntemi ölçü başına bilginin niceliğini etkiler. Örnekleme ile örneklem hacmi birlikte bir örneklemdeki gerekli bilginin toplam miktarını kontrol eder.

Araştırmacı örneklem hacminin seçim problemiyle karşılaşmadan önce bir deney planlamasında biraz gelişme yapar. Gerçekten belki de istatistikçinin kendisine en sık sorulan sorudan biri : “örneklem ne kadar ölçüm içermelidir?”. Ne yazık ki deneycinin ne kadar bilgi almayı istediğini bilmeden istatistikçi bu soruyu cevaplayamaz. Mutlaka örneklemdeki bilginin toplam miktarını değerlendirme, metodun iyilik ölçüsünü etkileyecek ve deneyci tarafından açıkça belirtilmelidir. Tahminin açıkça belirtilmesi hatırlanarak deneycinin tahminini ne kadar doğrulukta yapmak istediği bilinmelidir. Bu doğruluk (ya da kesinlik) tahmin hatası üzerine bir sınırlama ile açık olarak ifade edilebilir.

Örnek vermek gerekirse; bir kimyasal ürünün günlük ortalama üretimi μ tahmin edilmek istensin. Ayrıca bu tahmin %95 olasılıkla 5 tondan daha küçük bir tahmin hatasıyla gerçekleştirilmek istensin. Öyle ise $\frac{2\sigma}{\sqrt{n}} = 5$ ton yazılabilir. Her iki tarafın karesi alınır ve gerekli işlemler yapılırsa, $n = \frac{4\sigma^2}{25}$ eşitliği elde edilir. Buradan görüldüğü gibi kitle standart sapması (veya kitle varyansı) bilinmediği sürece n’in sayısal değeri elde edilemez. Zira $\bar{Y}$’nın değişkenliği kitleden çekilen örneklemin değişkenliğine bağlıdır.

σ’nın gerçek değerinin olmaması durumunda onun örneklemden hesaplanan yansız tahmini kullanılabilir. Eğer bu örnekte $s = 21$ ton olarak hesaplanmış ise $n = \frac{4\sigma^2}{25} = \frac{4(21)^2}{25} = 70.56 \cong 71$ olarak bulunur.

Burada açıklanan örneklem hacminin seçim metodu Tablo_I.1’de verilen büyük hacimli tahmin işlemleri için geçerlidir. Deneyci $(1-\alpha)$ güvenilirlik seviyesinde tahmin hatası için arzu ettiği sınırı kesin olarak belirlemelidir. Örneğin, eğer θ bir parametre ve b de arzu edilen (kabul edilebilir) hata sınırı ise $b = z_{\frac{\alpha}{2}}\sigma_{\hat{\theta}}$ eşitliği yazılabilir.

Örnek-I.9: Bir psikolojik deneyde bireyler *A* ve *B* gibi iki farklı tepkime ile cevap vermektedir. Böylesi bir çalışmada araştırmacı eğer bir birey A biçiminde bir tepki göstermişse π oranını tahmin etmek istediği düşünülsün. Bu durumda araştırmacı deneye kaç birey almalıdır. Araştırmacının %90 güvenle %4’den daha küçük tahmin hatasını kabul edebileceği de düşünülsün. Ayrıca araştırmacı π’nin de ğerçek değerinin 0.60 olacağını beklemekte olsun.

Çözüm-I.9: Öncelikle güvenilirlik seviyesi, $1-\alpha = 0.90$ ve böylece $\alpha = 0.10$ olup $\frac{\alpha}{2} = 0.05$ $=> z_{\frac{\alpha}{2}} = z_{0.05} = 1.645$ olduğuna göre $z_{\frac{\alpha}{2}}\sigma_P = 0.04$ veya $1.645\sqrt{\frac{0.6(0.4)}{n}} = 0.04$ ifadesinde gerekli işlemler yapılırsa $n \cong 406$ sonucu elde edilir.

- ✓ **Not-I.4:** Eğer deneyci (araştırmacı) $p = 0.60$ olarak ümit etmemiş olsaydı eşit olasılıklı (yani iki durumlu) deneylerde olduğu gibi $p = 0.50$ değerini kullanabilirdi. Bu durumda ise mümkün en büyük örneklem hacmi ise $n \cong 423$ olarak bulunur.

Örnek-I.10: Bir araştırmacı belli bir montajlamada çalışan sanayi işçilerinin performansına ilişkin iki farklı çalışma yönteminin etkisini karşılaştırmak istemektedir. Seçilen işçiler eşit örneklem hacimleriyle iki gruba ayrılmıştır. Birinci grup işçilerine çalışma yöntemi *A*, ikincisine ise yöntem *B* uygulanmıştır. Her bir işçiye montajlama yaptırılmış ve montajlamayı tamamlama süreleri (dk olarak) kaydedilmiştir. Her iki grubun da ölçümlerinin açıklığının (değişkenliğinin) aynı ve yaklaşık olarak 8 dk olacağı beklenmektedir. Eğer iki işçi grubunun ortalama montajlama süreleri arasındaki farka ilişkin tahmini %95 güvenilirlikle 1 dk içinde doğru kabul edilirse her bir çalışma grubuna kaçar adet işçi alınmalıdır.

Çözüm-I.10: $2\sigma_{(\bar{Y}_1-\bar{Y}_2)} = 1\text{dk}$ ve $n_1 = n_2 = n$ olduğuna göre $2\sqrt{\frac{(\sigma_1^2+\sigma_2^2)}{n}} = 1\text{dk}$ yazılabilir. Açıklık yaklaşık 4σ'ya eşittir. Öyle ise $4\sigma \approx 8 => \sigma \approx 2$ yazılabilir. He bir montajlama yönteminin değişkenliği aynı olduğuna göre $\sigma_1^2 = \sigma_2^2 = \sigma^2$ yazılabilir. Böylece $2\sqrt{\frac{2\sigma^2}{n}} = 1$ dk olarak elde edilir. Sonuç olarak; son eşitlikten $n = 32$ bulunur. Bu ise her bir işçi grubundan örnekleme alıncak işçi sayısıdır. Yani toplam 64 işçi iki gruba eşit olarak tesadüfen dağıtılacaklardır.

UYGULAMA-V

Uygulama-V.1: Y, p parametreli bir binom tesadüfi değişkeni olsun. Takipte verilen durumlar altında %95 güvenilirlik ve %5 hata içinde p'nin tahmini için gerekli örneklem hacmini hesaplayınız. a) Eğer p değeri, önceki çalışmalardan yaklaşık olarak 0.90 ise ve b) Eğer p hakkında hiçbir bilgi yoksa.

Çözüm-V.1: a) $2\sqrt{\frac{pq}{n}} = 0.05 => 2\sqrt{\frac{0.9(0.1)}{n}} = 0.05 => n = \frac{0.36}{0.0025} = 144$ bulunur. b) p hakkında hiçbir bilgi yoksa en yüksek örneklem hacmini veren oran tercih edilmelidir. Bu ise p'nin 0.50 alınması anlamına gelir. Böylece $2\sqrt{\frac{0.5(0.5)}{n}} = 0.05 => n = \frac{1.00}{0.0025} = 400$ bulunur.

Uygulama-V.2: bir avcılık kulübü verilen bir resmi avlanma sezonunda lisanslı bütün avcılardan her bir avcının avlandığı ortalama gün sayısını tahmin etmek istemektedir. Tahmin hatası 2 avlanma gününe eşitse araştırmaya kaç avcı alınmalıdır. Önceki çalışmalardan kitle standart sapmanın yaklaşık 10 gün olduğu bilinmektedir.

Çözüm-V.2: $2\sigma_{\bar{Y}} = 2$ gün $=> n = \sqrt{\frac{10^2}{n}} = 2$ gün $=> n = \frac{400}{4} = 100$ avcı yeterlidir.

I.8 μ ve $(\mu_1 - \mu_2)$ için küçük örneklem güven aralığı

Bu bölümde tartışılacak güven aralıkları normal dağılımdan seçilen örneklemler içindir. Burada tartışılacak güven aralığı her türlü örneklem hacimleri için uygundur ve kitle normal olmasa bile tatminkar sonuçlar vermektedir.

$Y_1, Y_2, \cdots, Y_n$ normal kitleden seçilmiş tesadüfi bir örneklemin örneklem ortalaması $\bar{Y}$ ve örneklem varyansı da S^2 olsun. Kitle varyansı bilinmediğinde ve örneklem hacmi de çok küçük olduğunda kitle ortalaması için güven aralığı pekâlâ oluşturulabilir. Bunu başarmak için $T = \frac{(\bar{Y}-\mu)}{S/\sqrt{n}}$ dağılımından yararlanılır ki $(n-1)$ serbestlik dereceli (sd) t dağılımı adını alır. Bu T dönüşüm değişkeni μ'nün bir güven aralığının oluşturulması için döngüsel niceliğe sahiptir. t için standart normal dağılımda olduğu gibi hazır tablolar geliştirilmiştir. Buradan $P\left(-t_{\alpha/2} \leq T \leq t_{\alpha/2}\right) = 1 - \alpha$ eşitliği yazılabilir. t dağılımı uç noktaların daha yaygın olması dışında standart normal ($z \sim N(0,1)$) dağılımın olasılık yoğunluk fonksiyonuna oldukça benzerdir. $t_{\alpha/2}$'nin tablo değeri hem sd $(n-1)$'e ve hem de güvenilirlik seviyesi $(1-\alpha)$ değerine ya da birinci tip hata olasılığı α'ya bağlıdır. Bu durumda μ'nün güven aralığı, $\bar{Y} \pm t_{\alpha/2}\left(S/\sqrt{n}\right)$ formundadır. Burada $\bar{Y} - t_{\alpha/2}\left(S/\sqrt{n}\right)$ alt güven limiti ve $\bar{Y} + t_{\alpha/2}\left(S/\sqrt{n}\right)$ üst güven limitini göstermektedir.

Örnek-I.11: Bir mermi üreticisi, yeni bir barutla ürettiği mermileri 8 mermi ile test etmiştir. Gözlenen değerler saniyede metre olarak sırasıyla: 917; 892; 895; 904; 914; 917; 896 ve 886. %95 güvenle bu tip mermilerin saniyede gerçek ortalama hızı μ için güven aralığı tahmin ediniz.

Çözüm-I.11: Mermi hızları Y_i ile gösterilsin ve normal dağıldığı varsayılsın. μ için güven aralığı, $\bar{Y} - t_{\alpha/2}\left(S/\sqrt{n}\right)$ tahmin edicisi kullanılarak tahmin edilebilir. Bu örneklem için $\bar{y} = 903$ m, $s = 12.2$ metredir. t dağılım tablosundan $n-1 = 8-1 = 7$ sd ve $\frac{\alpha}{2} = 0.025$ önem seviyesinde $t_{7;0.025} = 2.365$ bulunur. Böylece $\bar{Y} - t_{\frac{\alpha}{2}}\left(\frac{S}{\sqrt{n}}\right) = 903 \pm 2.365\left(\frac{12.2}{\sqrt{8}}\right) = 903 \pm 10.2$ elde edilir. Öyle ise μ'nün %95 güven aralığı, $(892.8, 913.2)$ yazılabilir.

Şimdi de iki normal kitleden ortalamaların karşılaştırılması düşünülsün. Birinci kitlenin ortalaması μ_1 ve varyansı σ_1^2; aynı şekilde ikinci kitlenin ortalaması μ_2 ve varyansı σ_2^2 ile gösterilsin. Eğer $\sigma_1^2 = \sigma_1^2 = \sigma^2$ kabul edilirse T tesadüfi değişkenine dayanarak $(\mu_1 - \mu_2)$ için bir güven aralığı oluşturulabilir.

Eğer iki bağımsız örneklemden elde edilen örneklem ortalamalarının tahmin edicileri sırası ile $\bar{Y}_1$ ve $\bar{Y}_2$ ise $(\mu_1 - \mu_2)$ için büyük örneklem güven aralığı, $Z = [(\bar{Y}_1 - \bar{Y}_2) - (\mu_1 - \mu_2)]\left[\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right]^{-1}$ tesadüfi değişkeni kullanılarak geliştirilebilir. Bilindiği gibi bu değişken bir standart normal dağılıma sahiptir. $\sigma_1^2 = \sigma_1^2 = \sigma^2$ varsayımı altında bu eşitlik $Z = [(\bar{Y}_1 - \bar{Y}_2) - (\mu_1 - \mu_2)]\left[\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}\right]^{-1}$ şeklinde yazılabilir. Bir t dağılımı niceliğini kullanmak için σ^2'nin ortak bir tahmin edicisi bulunmalıdır.

$Y_{11}, Y_{12}, \cdots, Y_{1n_1}$ birinci kitleden tesadüfi örneklem ve $Y_{21}, Y_{22}, \cdots, Y_{2n_2}$ de ikinci kitleden aynı özelliğe sahip tesadüfi örneklem olsun. Öyle ise bu iki örneklemin ortalaması $\bar{Y}_1 = \frac{1}{n_1}\sum_{i=1}^{n_1} Y_{1i}$ ve $\bar{Y}_2 = \frac{1}{n_2}\sum_{i=1}^{n_1} Y_{2i}$ tahmin edicileri ile bulunur. Örneklem verilerinden σ^2'nin ortak yansız tahmin edicisi ise $S^2 = \left\{\left[\sum_{i=1}^{n_1}(Y_{1i} - \bar{Y}_1)^2\right] + \left[\sum_{i=1}^{n_2}(Y_{2i} - \bar{Y}_2)^2\right]\right\}(n_1 + n_2 - 2)^{-1} = [(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2](n_1 + n_2 - 2)^{-1}$ ile verilir. Bu son eşitlikte S_j^2, $j = 1{,}2$ olmak üzere j'inci örneklemin varyans tahmin edicisidir. Burada $\frac{[(n_1+n_2-2)S^2]}{\sigma^2} \sim \chi^2_{n_1+n_2-2}$olup $(n_1 - 1)$ ve $(n_2 - 1)$ serbestlik dereceli iki bağımsız ki-kare tesadüfi değişkenin toplamıdır. Böylece $\frac{[(n_1+n_2-2)S^2]}{\sigma^2}$, $v = (n_1 + n_2 - 2)$ sd'li bir ki-kare dağılımına sahiptir. Buradan χ^2 ve Z değişkenleri döngüsel nicelik oluşturmak için kullanılırsa $T = Z(\chi^2/v)^{-1} = [(\bar{Y}_1 - \bar{Y}_2) - (\mu_1 - \mu_2)]\left[S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}\right]^{-1}$ elde edilir. Bu sonuç ise $v = (n_1 + n_2 - 2)$ sd'li bir t dağılımıdır.

Bu sonuçtan hareketle $(\mu_1 - \mu_2)$ için güven aralığı tahmin edicisi, $(\bar{Y}_1 - \bar{Y}_2) \pm t_{\alpha/2} S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$ formunda verilebilir. Bu ifadede $t_{\alpha/2}$, $\alpha/2$ ile $v = (n_1 + n_2 - 2)$ ile birleştirilmiş t dağılım tablo değeridir.

Örnek-I.12: Bir işletme atölyesinde montajlama işleminde istihdam edilecek yeni işçilerin işi tam kavrayabilmeleri için yaklaşık olarak bir ay eğitim görmeleri gerekmektedir. Eğitim için yeni bir yöntem geliştirilmiş ve bu yöntem eski geleneksel yöntemle karşılaştırılmak istenmiştir. 18 yeni işçi eşit sayıda ve tesadüfen iki gruba ayrılmış ve bir ay eğitime tabi tutulmuştur. Eğitim sonunda her bir işçinin montajlama süresi dakika olarak Tablo-I.3'de verildiği gibi elde edilmiştir. Montajlama sürelerinin normal dağılımlı ve her iki yöntem için varyansların yaklaşık olarak eşit kabulü altında gerçek ortalama farkı $(\mu_1 - \mu_2)$ için %95 güven aralığı tahmin ediniz.

Tabo-I.3: Örnek-I.12 İçin Veriler.

Yöntem	Montajlama Süresi (dk)
Standart	32; 37; 35; 28; 41; 44; 35; 31; 34
Yeni	35; 31; 29; 25; 34; 40; 27; 32; 31

Çözüm-I.12: $\bar{y}_1 = 35.22$ dk; $\bar{y}_2 = 31.56$ dk ; $\sum_{i=1}^{9}(y_{1i} - \bar{y}_1)^2 = 195.56$ ve $\sum_{i=1}^{9}(y_{2i} - \bar{y}_2)^2 = 160.22$ olup ortak varyans, $s^2_{\bar{y}_1-\bar{y}_2} = \frac{(195.56+160.22)}{(9+9-2)} = 22.24$ ve böylece $s_{\bar{y}_1-\bar{y}_2} = 4.71$ dk bulunur. Ayrıca $t_{16;0.025} = 2.120$ (t tablosundan) bulunur. Böylece istenen güven aralığı, $(\bar{Y}_1 - \bar{Y}_2) \pm t_{\alpha/2}S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$ tahmin edicisi kullanılarak $(35.22 - 31.56) \pm 2.12(4.71)\sqrt{\frac{1}{9} + \frac{1}{9}}$ $=> 3.66 \pm 2.22$ olarak elde edilir. O halde bu güven aralığı $(1.44, 5.88)$ formunda da yazılabilir. Böylece bu güven aralığından da görülür ki aralık sıfırı içermemektedir. Yani $\mu_1 > \mu_2$ olduğundan standart yöntemin montajlama süresi daha büyük olup yeni yöntemin montajlamada daha iyi olduğu %95 güvenilirlikle söylenebilir.

✓ **Not-I.5:** Örneklem hacmi *n* artarken *t*-dağılımının serbestlik derecesi de buna bağlı olarak artacak ve *t*-dağılımı standart normal (*z* dağılımı) dağılıma yaklaşacaktır. Böylece küçük örneklemli güven aralıkları büyük örneklemli güven aralıklarına eşit olacaktır. Sonuç olarak, *t*'nin sd 30'dan büyük olduğunda güven aralıkları hemen hemen büyük örneklemli güven aralıkları ile aynı olur.

UYGULAMA-VI

Uygulama-VI.1: 14 seçkin bayan voleybol oyuncusunun vücut yağ ölçümlerinin ortalama ve standart sapması %'de olarak sırasıyla 17.9 ve 3.6'dır. Bu 14 oyuncunun bütün seçkin bayan voleybolcular arasından seçilmiş bir örneklem olduğunu kabul ederek bayan voleybolular bayan voleybolcular kitlesi yağ oranı ortalaması için %95 güven aralığı tahmin ediniz.

Çözüm-VI.1: $\bar{y} = \%17.9$, $s = \%3.6$ ve $n = 14$ olduğuna göre $t_{\alpha/2} = t_{0.025} = 2.160$ (t tablosundan) $\bar{Y} \pm t_{\alpha/2}(s/\sqrt{n})$ tahmin edicisi yardımı ile $0.179 \pm 2.16(0.036/\sqrt{14})$ $=>$ $0.179 \pm 0.021 =>$ $(0.158, 0.200)$ bulunur.

Uygulama-VI.2: Mezuniyet sınavına giren 16 kişilik bir tesadüfi örneklemin puanlar ortalaması 540 puan ve standart sapması 50 puan olarak hesaplanmıştır. Puanların normal dağılımlı olduğu varsayımı altında mezuniyet sınavına katılan kişilerin mezuniyet kitle ortalaması için %95 güven aralığı tahmin ediniz.

Çözüm-VI.2: $\bar{y} = 540$, $s = 50$ ve $n = 16$ olduğuna göre $t_{\alpha/2} = t_{0.025} = 2.131$ (t tablosundan) $\bar{Y} \pm t_{\alpha/2}(s/\sqrt{n})$ tahmin edicisi yardımı ile $540 \pm 2.131(50/\sqrt{16}) => 540 \pm 26.64 =>$ $(513.36, 566.64)$ bulunur.

Uygulama-VI.3: Koşucular için oksijen tüketim oranı fizyolojik olarak önemlidir. Oksijen tüketim oranları arasındaki farkı ortaya çıkarmak amacıyla bir spor akademisindeki öğrencilere iki farklı koşu yöntemi uygulanmıştır. Bir gruba sürekli koşu, diğer gruba da kesikli koşu yöntemi uygulanmıştır. Çalışmaya ilişkin bazı veri ve hesaplamalar Tablo-I.4'de verilmiştir. Ölçüm değerlerinin normal dağılımlı ve varyanslarının da eşit olduğu varsayımı altında kitle ortalamaları arasındaki farka ilişkin %95 güven aralığı tahmin ediniz.

Tabo-I.4: Uygulama-VI.3 İçin Veriler.

Yöntem	n_j	$\bar{y}_j$	s_j^2
Sürekli	9	43.71	5.88
Kesikli	7	39.63	7.68

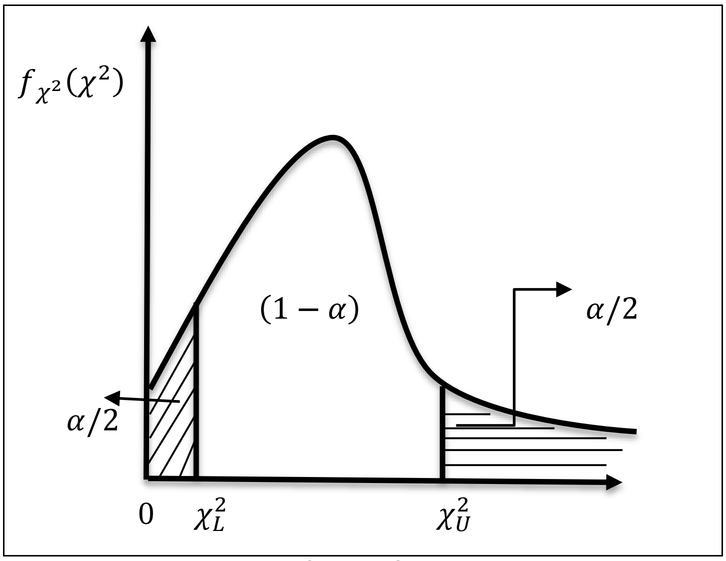
Çözüm-VI.3: $\sigma_1^2 = \sigma_2^2 = \sigma^2$ olduğuna göre $s^2 = [(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2]/(n_1 + n_2 - 2)$ eşitliği kullanılarak $s^2 = 6.651$ bulunur. $v = n_1 + n_2 - 2 = 9 + 7 - 2 = 14$ sd 0.025 önem seviyesinde $t_{0.025} = 2.145$ (t tablosundan) $(\bar{Y}_1 - \bar{Y}_2) \pm t_{\alpha/2} S \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$ tahmin edicisi kullanılarak %95 güven aralığı tahmin edilebilir. $(43.71 - 39.63) \pm 2.145(1.3) => 4.08 \pm 2.789 => (1.292, 6.869)$ olarak bulunur. Bu sonuca göre oluşturulan güven aralığı sıfırı içermediğinde sürekli koşucuların oksijen kayıp oran ortalaması kesikli koşuculardan daha yüksek olduğu söylenebilir.

Uygulama-VI.4: Bir fabrikada aynı iş için A ve B gibi iki tip makine çalışmaktadır. A tip makinenin haftalık bakım maliyeti Y, μ_1 ortalama ve σ^2 varyans ile normal dağılıma sahiptir. B tip makinenin haftalık bakım maliyeti X, normal dağılımlı; fakat μ_2 ortalama ve $3\sigma^2$ varyanslıdır. Fabrika için her bir haftalık beklenen bakım masrafı, $2\mu_1 + \mu_2$ olduğuna göre eğer $Y_1, \cdots, Y_n$, A makinesinden ve $X_1, \cdots, X_m$ de B makinesinden tesadüfi ve bağımsız iki örneklem ise fabrikanın haftalık bakım masrafı, $(2\mu_1 + \mu_2)$ için $(1 - \alpha)$ önem seviyesinde güven aralığı oluşturunuz. **Not:** $n + m < 30$, Y ile X bağımsız ve ayrıca kitle varyansı σ^2 de bilinmemektedir.

Çözüm-VI.4: μ_1'in yansız tahmin edicisi, $\bar{Y}$ ve μ_2'nin yansız tahmin edicisi, $\bar{X}$ olup haftalık bakım maliyeti $(2\mu_1 + \mu_2)$'nün yansız tahmin edicisi, $Z = 2\bar{Y} + \bar{X}$ yazılabilir. $V(Z) = 4V(\bar{Y}) + V(\bar{X})$ olduğuna göre $V(\bar{Y}) = \sigma^2/n$ olur. Benzer şekilde $V(\bar{X}) = 3\sigma^2/m$ bulunur. Böylece $V(Z) = 4(\sigma^2/n) + 3\sigma^2/m$ yazılabilir. σ^2 bilinmediğine göre onun yansız tahmin edicisi S^2 tanımlanmalıdır. Öyle ise $S_1^2 = \frac{1}{(n-1)} \sum_{i=1}^{n} (Y_i - \bar{Y})^2$ ve $S_2^2 = \frac{1}{(m-1)} \sum_{i=1}^{m} (X_i - \bar{X})^2$ ile tanımlanırsa $S^2 = \left[(n-1)S_1^2 + \frac{(m-1)}{3} S_2^2\right]/(n + m - 2)$ => $s = \sqrt{s^2}$ yazılır. Böylece fabrikanın haftalık bakım masrafı için güven aralığı tahmin edicisi, $Z \pm t_{\alpha/2} s \sqrt{\frac{4}{n} + \frac{3}{m}}$ elde edilir. Burada $t_{\alpha/2}$, $v = (n + m - 2)$ sd'de ve $\alpha/2$ önem seviyesinde t-dağılımının tablo değeridir.

I.9 Kitle Varyansı σ^2) İçin Güven Aralığı

Araştırmacı çoğu kez σ^2'yi bilemez ve onu tahmin etmeyi ister. Bilindiği gibi $S^2 = (n-1)^{-1}\sum_{i=1}^{n}(Y_i - \bar{Y})^2$, σ^2'nin yansız bir tahmin edicisidir. σ^2 için aralık tahmin işlemini uygulamak için döngüsel nicelik bulunmalıdır. Şimdi $Y_1, \cdots, Y_n$ bilinmeyen bir ortalama μ ve bilinmeyen bir varyans σ^2 olan normal dağılımdan bir örneklem olsun. $\sigma^{-2}\sum_{i=1}^{n}(Y_i - \bar{Y})^2 = (n-1)S^2/\sigma^2$, $(n-1)$ sd'li bir ki-kare $(\chi^2_{(n-1)})$ dağılımına sahiptir. Böylece σ^2 için bir döngüsel metot geliştirilebilir. Yani herhangi bir güvenilirlik seviyesi için $P(\chi^2_L \leq [(n-1)S^2/\sigma^2] \leq \chi^2_U) = 1 - \alpha$ yazılabilir. Burada indislerdeki L ve U sırasıyla ki kare olasılık yoğunluk fonksiyonunun alt ve üst limit değerleridir. Önceden olduğu gibi amaç $(1-\alpha)$ güvenle en dar güven aralığı oluşturmaktır; ancak genellikle bunu başarmak zordur. Bir ki-kare olasılık yoğunluk fonksiyonu kritik bölgelerin gösterimi Şekil-I.6'da verilmiştir.



Şekil-I.6: χ^2_L ve χ^2_U Değerlerin Yeri.

Gerekli düzenlemeler yapıldıktan sonra σ^2 için $(1-\alpha)$ güven aralığı, $P\left\{\frac{(n-1)S^2}{\chi^2_U} \leq \sigma^2 \leq \frac{(n-1)S^2}{\chi^2_L}\right\} = (1-\alpha)$ ile verilir. Buradan σ^2'nin güven aralığı $\left(\frac{(n-1)S^2}{\chi^2_U}, \frac{(n-1)S^2}{\chi^2_L}\right)$ ile ifade edilir.

Örnek-I.13: Bir araştırmacı her hangi bir işitme kaynağının değerini ölçmek için planlanan aletin değişimini kontrol etmek amacıyla bu aletle aynı şartlar altında üç adet ölçüm gerçekleştirmiş ve 4.1; 5.2 ve 10.2 değerlerini elde etmiştir. %90 güvenilirlikle σ^2 için bir güven aralığı tahmin ediniz.

Çözüm-I.13: Ölçüm değerlerinin normal dağılımlı olduğu varsayımı altında $s^2 = 10.57$ bulunur. $\alpha/2 = 0.05$ ve $(n-1) = (3-1) = 2$ sd'nde $\chi^2_L = \chi^2_{2;0.05} = 0.103$ ve $\chi^2_U = \chi^2_{2;0.95} = 5.991$ (ki-kare tablosundan) bulunur. Böylece σ^2 için %90 güven aralığı ise $\left(\frac{(n-1)S^2}{\chi^2_U}, \frac{(n-1)S^2}{\chi^2_L}\right) => \left(\frac{2(10.57)}{5.991}, \frac{2(10.57)}{0.103}\right) => (3.53, 205.24)$ bulunur. Örneklem hacmi küçük olduğundan σ^2'nin bu örnekte elde edilen güven aralığı çok geniş bulunmuştur.

I.10 Uygulama-VII

Uygulama-VII.1: Bir elektronik fabrikasında çalışan çok sayıda mühendis grubundan tesadüfi olarak 21 tanesi seçilmiştir. Haftalık çalışma sürelerine ilişkin satndart sapma 7 saat olarak hesaplanmıştır. Ölçümlerin normal dağılıma sahip olduğu varsayımı altında fabrikada çalışan tüm mühendisler kitlesi ait haftalık çalışma sürelerinin kitle varyansı için %90 güvenilirlikle bir aralık tahmin ediniz.

Çözüm-VII.1: $s = 7 \ => \ s^2 = 49$ ve $\alpha/2 = 0.05$ ve $(n-1) = (21-1) = 20$ sd'nde $\chi_L^2 = \chi_{2;0.05}^2 = 10.8508$ ve $\chi_U^2 = \chi_{2;0.95}^2 = 31.4104$ (ki-kare tablosundan) bulunur. . Böylece σ^2 için %90 güven aralığı ise $\left(\frac{(n-1)S^2}{\chi_U^2}, \frac{(n-1)S^2}{\chi_L^2}\right) => \left(\frac{20(49)}{31.4104}, \frac{20(49)}{10.8508}\right) => (31.20, 90.92\ )$ bulunur.

Uygulama-VII.2: Bir üniversitede görev yapan beş profesörlük tesadüfi örnekleme ilişkin yaşlar sırasıyla: 39; 54; 61; 72 ve 59 olsun. Bu bilgiyi kullanarak ilgili üniversitede görev yapan tüm profesörlerin yaşlarının varyansı için %99 güvenilirlikle bir aralık tahmin ediniz. Yaşların normal dağılıma sahip olduğu varsayınız.

Çözüm-VII.2: Verilerden $s^2 = 144.5$ bulunur. $\alpha/2 = 0.005$ ve $(n-1) = (5-1) = 4$ sd'nde $\chi_L^2 = \chi_{4;0.005}^2 = 0.20699$ ve $\chi_U^2 = \chi_{4;0.995}^2 = 14.8602$ (ki-kare tablosundan) bulunur. Böylece σ^2 için %99 güven aralığı ise $\left(\frac{(n-1)S^2}{\chi_U^2}, \frac{(n-1)S^2}{\chi_L^2}\right) => \left(\frac{4(144.5)}{14.8602}, \frac{4(144.5)}{0.20699}\right) =>$ $(38.90, 2792.41\ )$ bulunur.

I.11 Biribci Bölümle İlgili Çalışma Soru ve Çözümleri

ÇS–1: Olasılık yoğunluk fonksiyonu, $f(y) = \beta^{-1}e^{-y/\beta}, y > 0$ olan bir üstel dağılımın tesadüfî bir örnek değeri, $Y = 35$ olsun. $U = Y/\beta$ dönüşümünü kullanarak β için %90 güven aralığı tahmin ediniz.

ÇSÇ-1: U'nun o.y.f., $f_U(u) = e^{-u}, u > 0$ olup standart (birim) üstel dağılıma sahiptir. U için % 90 güven aralığı: P{a$\leq U \leq$b}=0.90 $\Rightarrow$ P{0.051$\leq U \leq$2.996}=0.90 $\Rightarrow$ P{0.051$\leq \frac{Y}{\beta} \leq$2.996}=0.90 $\Rightarrow$ P{$\frac{Y}{2.996} \leq \beta \leq \frac{Y}{0.051}$}=0.90 Y=35 $\Rightarrow$ P{$\frac{35}{2.996} \leq \beta \leq \frac{35}{0.051}$}=0.90 $\Rightarrow$ P{11.682$\leq \beta \leq$686.275}=0.90 bulunur.

ÇS–2: $Y_1 \sim N(\mu_1, \sigma_1^2)$ ve $Y_2 \sim N(\mu_2, \sigma_2^2)$ iki ayrı bağımsız normal kitleden tesadüfî olarak çekilen 30'ar birimlik örneklemlerin ortalaması sırası ile $\bar{y}_1 = 18$ ve $\bar{y}_2 = 12$; varyansları da sırası ile $s_1^2 = 16$ ve $s_2^2 = 25$ olarak hesaplanmış ise $(\mu_1 - \mu_2)$ parametresi için %95 güven aralığı tahmin ediniz ve sonucu yorumlayınız. (NOT: Tablo değerini 2 alınız.)

ÇSÇ-2: $z_T = 2$ $\hat{\sigma}_{\bar{Y}_1 - \bar{Y}_2} = \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$ $s_{\bar{Y}_1 - \bar{Y}_2} = \sqrt{\frac{16}{30} + \frac{25}{30}} = 1.169$ bulunur. Verilen değerler $(\bar{Y}_1 - \bar{Y}_2) \pm z_{\frac{\alpha}{2}}\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$ ifadesinde değerlendirilirse $6 \pm 2(1.169) = 6 \pm 2.338$ olup $P(3.662 \leq (\mu_1 - \mu_2) \leq 8.338) = 0.95$ bulunur. %95 güvenilirlikle tahmin edilen aralık sıfırı içermediğinden birinci kitlenin ortalamasının ikinci kitlenin ortalamasından daha büyük olduğu söylenebilir.

ÇS-3: Ortalaması μ olan bir Poisson dağılımın 9 birimlik tesadüfî bir örneklemi, $Y_1,\cdots,Y_9$ olsun. $\hat{\mu}=c[Y_1(Y_1-Y_2)+\cdots+Y_8(Y_8-Y_9)]$ Tahmin edicisi tanımlandığında $\hat{\mu}$'yı μ için yansız yapacak c sabitini bulunuz.

ÇS-4: Olasılık yoğunluk fonksiyonu, $f(y)=\beta^{-1}e^{-y/\beta},y>0$ olan bir üstel dağılımın tesadüfî bir örnek değeri, $y=14$ olsun. $U=Y/\beta$ dönüşümünü kullanarak β için %95 güven aralığı tahmin ediniz.

ÇS–5: $Y_1{\sim}B(n_1,\pi_1)$ ve $Y_2{\sim}B(n_2,\pi_2)$ iki ayrı binom kitlesinden $n_1=n_2=100$'er birimlik örneklemlere ilişkin değişken değerleri $y_1=75$ ve $y_2=90$ olarak gözlemlenmiştir. $(\pi_1-\pi_2)$ parametresi için %95 güven aralığı tahmin ediniz ve sonucu yorumlayınız. (NOT: Tablo değerini 2 alınız.)

ÇS-6: Ortalaması λ olan bir Poisson dağılımın10 birimlik tesadüfî bir örneklemi, $Y_1,\cdots,Y_{10}$ olsun. $\hat{\lambda}=k[Y_1(Y_1-Y_2)+\cdots+Y_9(Y_9-Y_{10})]$ Tahmin edicisi tanımlandığında $\hat{\lambda}$'yı λ için yansız yapacak k sabitini bulunuz.

ÇS-7: Olasılık yoğunluk fonksiyonu, $f(y)=\beta^{-1}e^{-y/\beta},y>0$ olan bir üstel dağılımın tesadüfî bir örnek değeri, $y=13$ olsun. $U=Y/\beta$ dönüşümünü kullanarak β için %95 güven aralığı tahmin ediniz.

ÇS-8: $Y_1{\sim}B(n_1,\pi_1)$ ve $Y_2{\sim}B(n_2,\pi_2)$ iki ayrı binom kitlesinden $n_1=n_2=100$'er birimlik örneklemlere ilişkin değişken değerleri $y_1=70$ ve $y_2=85$ olarak gözlemlenmiştir. $(\pi_1-\pi_2)$ parametresi için %95 güven aralığı tahmin ediniz ve sonucu yorumlayınız. (NOT: Tablo değerini 2 alınız.)

ÇS-9: Ortalaması μ olan bir Poisson dağılımın n birimlik tesadüfî bir örneklemi, $Y_1,\cdots,Y_n$ olsun. $\hat{\mu}=c[Y_1(Y_1-Y_2)+\cdots+Y_{n-1}(Y_{n-1}-Y_n)]$ tahmin edicisi tanımlandığında $\hat{\mu}$'yı μ için yansız yapacak c sabitini bulunuz.

ÇS-10: Olasılık yoğunluk fonksiyonu, $f(y)=\beta^{-1}e^{-y/\beta},y>0$ olan bir üstel dağılımın tesadüfî bir örnek değeri, $Y=25$ olsun. $U=Y/\beta$ dönüşümünü kullanarak β için %95 güven aralığı tahmin ediniz.