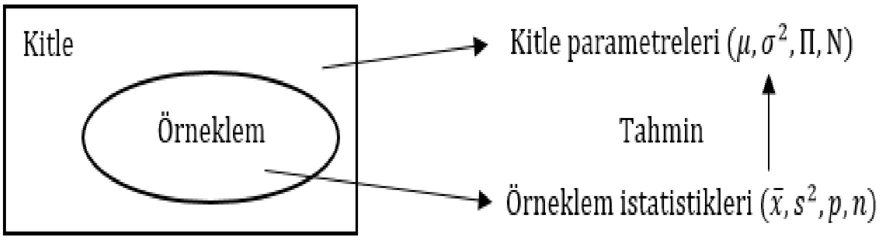
9. ÖRNEKLEME DAĞILIMI

Örneklem istatistiklerini kullanarak bilinmeyen kitle parametrelerini tahmin etme yöntemlerinin genel adına *çıkarımsal istatistik* denir.



Kitleden seçilen örneklemin istatistikleri ($\bar{x}$, s^2,p gibi) elde edilen her bir örneklem için değişkenlik göstereceğinden bu istatistikler birer rasgele değişkenlerdir ($\bar{X}$, S^2,P gibi) ve birer dağılıma sahiptirler.

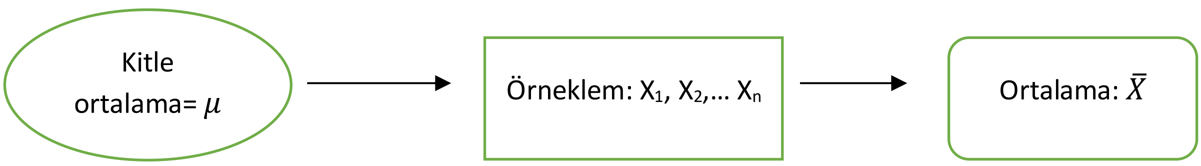
Tanım: N birimlik bir kitleden n büyüklüğünde çekilebilecek bütün örneklemler için hesaplanan herhangi bir istatistiğin dağılımına *örnekleme dağılımı* denir.

Tanım: μ ortalamalı σ^2 varyanslı N birimlik bir kitleden çekilebilecek (iadeli N^n ya da iadesiz $C(N,n)$ tane) bütün örneklemlerin ortalamasından oluşan $\bar{X}$ rasgele değişkeninin dağılımına *ortalamanın örnekleme dağılımı* denir.

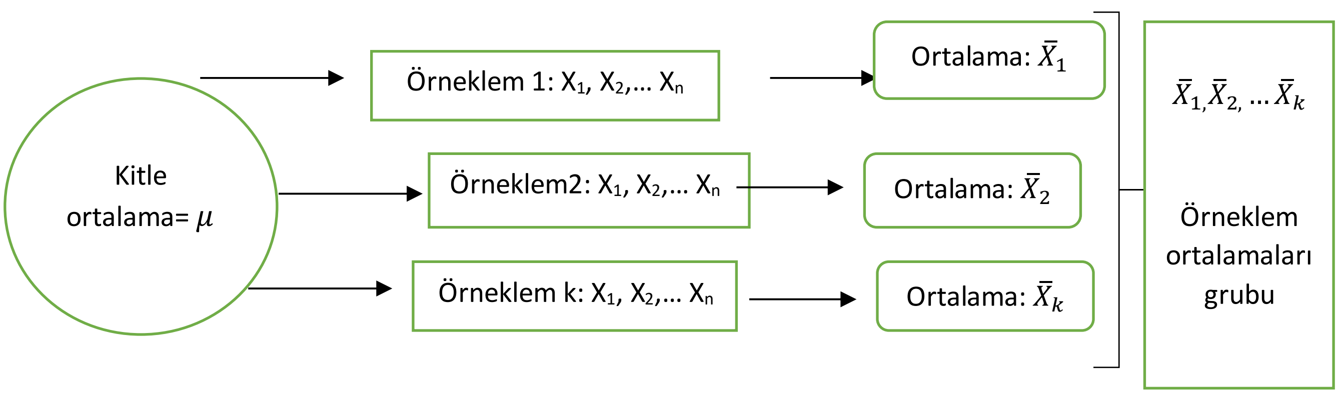
$E(\bar{X}) = \mu_{\bar{X}} = \mu$ olup iadeli örneklem için $V(\bar{X}) = \sigma_{\bar{X}}^2 = \frac{\sigma^2}{n}$ iadesiz örneklem için $V(\bar{X}) = \sigma_{\bar{X}}^2 = \frac{\sigma^2}{n} * \frac{N-n}{N-1}$ dir.

Not: Genel olarak biz kitledeki gözlem sayısını bilmeyiz ve sonsuz büyüklükte kabul ederiz. Bu nedenle $({N-n}/{N-1})$ değeri büyük N değerleri için 1'e yakınsayacağından iadeli ve iadesiz örneklemler için varyans $\sigma_{\bar{X}}^2 = \sigma^2/n$ olarak alınır ve standart sapması (hatası) da $\sigma_{\bar{X}} = \sigma/\sqrt{n}$ dir.

μ ortalamalı bir kitleden n birimlik bir örneklem seçilmesi



μ ortalamalı bir kitleden n birimlik k tane örneklem seçilmesi



Örnek: 2,4,6,8,10 olarak verilen kitle için yerine koymaksızın 2 birimlik örneklem seçiliyor. Örneklem ortalamasının beklenen değerini bulunuz.

C(5,2)=10 farklı örneklem söz konusudur. Bunları tabloda gösterelim:

Örneklem	Örneklem Ortalaması
2,4	**3**
2,6	**4**
2,8	**5**
2,10	**6**
4,6	**5**
4,8	**6**
4,10	**7**
6,8	**7**
6,10	**8**
8,10	**9**

Her bir sonucu 1/10 olasılıkla elde edebileceğimiz açıktır. Bu nedenle rasgele bir örneklem oluşmuş olur. Olasılıkları şu şekilde ifade de edelim:

P($\bar{X}$=9)=1/10, P($\bar{X}$=8)=1/10 P($\bar{X}$=7)=2/10, P($\bar{X}$=6)=2/10, P($\bar{X}$=5)=2/10 P($\bar{X}$=4)=1/10, P($\bar{X}$=3)=1/10.

$$E(\bar{X}) = \sum_i x_i f(x_i) = 3.0{,}1 + 4.0{,}1 + \cdots + 9.0{,}1 = 6$$

Diğer yandan $\mu = \frac{2+4+6+8+10}{5} = 6$ olup kitle ortalaması ile örneklem ortalaması çakışır.

Örnek: N=3 birimlik bir kitlenin elemanları $X_1 = 2, X_2 = 4 \; ve \; X_3 = 6$ olsun. Bu kitleden iadeli ve iadesiz olarak n=2 birimlik tüm mümkün örneklemler çekilsin. Örnekleme dağılımını oluşturup, dağılımın beklenen değeri (ortalaması) ve varyansını hesaplayınız.

Çözüm: Öncelikle kitle ortalaması ve varyansını bulalım.

$$\mu = \frac{\sum_{i=1}^{N} X_i}{N} = \frac{2+4+6}{3} = 4$$

$$\sigma^2 = \frac{\sum_{i=1}^{N} (X_i - \mu)^2}{N} = \frac{(2-4)^2 + (4-4)^2 + (6-4)^2}{3} = 2.67$$

a) *İadeli örneklemler*: 3 birimlik bir kitleden iadeli olarak alınan 2 birimlik örneklemlerin sayısı $N^n = 3^2 = 9$ dur. Bunlar

No	**Tüm mümkün örneklemler**	**Örneklem ortalamaları ($\bar{X}_i$)**
1	2 ve 2	2
2	2 ve 4	3
3	2 ve 6	4
4	4 ve 2	3
5	4 ve 4	4
6	4 ve 6	5
7	6 ve 2	4
8	6 ve 4	5
9	6 ve 6	6

Ortalamanın Örnekleme Dağılımı		
Örneklem ortalamaları ($\bar{X}$)	**Frekanslar (f)**	**Göreli frekanslar (olasılıklar $f(\bar{X})$)**
2	1	1/9
3	2	2/9
4	3	3/9
5	2	2/9
6	1	1/9
Toplam	**9**	**1**

Dağılımın beklenen değeri

$$E(\bar{X}) = \sum \bar{X} * f(\bar{X}) = 2 * \frac{1}{9} + 3 * \frac{2}{9} + 4 * \frac{3}{9} + 5 * \frac{2}{9} + 6 * \frac{1}{9} = \frac{36}{9} = 4$$
$$= \frac{\sum \bar{X} * f}{\sum f} = \frac{2 * 1 + 3 * 2 + 4 * 3 + 5 * 2 + 6 * 1}{9} = 4$$

olup $E(\bar{X}) = \mu = 4$

Dağılımın varyansı

$$V(\bar{X}) = \sum \left(\bar{X} - E(\bar{X})\right)^2 * f(\bar{X}) = (2-4)^2 * \frac{1}{9} + (3-4)^2 * \frac{2}{9} + \cdots + (6-4)^2 * \frac{1}{9} = 1.33$$
$$= \frac{\sum \left(\bar{X} - E(\bar{X})\right)^2 * f}{\sum f} = \frac{(2-4)^2 * 1 + (3-4)^2 * 2 + \cdots + (6-4)^2 * 1}{9} = 1.33$$

olup $V(\bar{X}) = \frac{\sigma^2}{n} = \frac{2.67}{2} = 1.33$

b) *İadesiz örneklemler*: 3 birimlik bir kitleden iadesiz olarak alınan 2 birimlik örneklemlerin sayısı $C(3,2) = 3$ dur. Bunlar

No	Tüm mümkün örneklemler	Örneklem ortalamaları ($\bar{X}_i$)
1	2 ve 4	3
2	2 ve 6	4
3	4 ve 6	5

Ortalamanın Örnekleme Dağılımı		
Örneklem ortalamaları ($\bar{X}$)	Frekanslar (f)	Göreli frekanslar (olasılıklar $f(\bar{X})$)
3	1	1/3
4	1	1/3
5	1	1/3
Toplam	3	1

Dağılımın beklenen değeri

$$E(\bar{X}) = \sum \bar{X} * f(\bar{X}) = 3 * \frac{1}{3} + 4 * \frac{1}{3} + 5 * \frac{1}{3} = \frac{12}{3} = 4$$

$$= \frac{\sum \bar{X} * f}{\sum f} = \frac{3 * 1 + 4 * 1 + 5 * 1 +}{3} = 4$$

olup $E(\bar{X}) = \mu = 4$

Dağılımın varyansı

$$V(\bar{X}) = \sum \left(\bar{X} - E(\bar{X})\right)^2 * f(\bar{X}) = (3-4)^2 * \frac{1}{3} + (4-4)^2 * \frac{1}{3} + (5-4)^2 * \frac{1}{3} = 0.67$$

$$= \frac{\sum \left(\bar{X} - E(\bar{X})\right)^2 * f}{\sum f} = \frac{(3-4)^2 * 1 + (4-4)^2 * 1 + (5-4)^2 * 1}{3} = 0.67$$

olup $V(\bar{X}) = \frac{\sigma^2}{n} * \frac{N-n}{N-1} = \frac{2.67}{2} * \frac{3-2}{3-1} = 0.67$

Tanım: Bir kitle parametresi tahmin edilirken kitle yerine örneklemin kullanılmasıyla oluşan hataya yani $\bar{X} - \mu$ farkına örnekleme hatası denir.

Teorem: μ ortalamalı σ^2 varyanslı normal dağılıma sahip bir kitleden çekilen örneklemin ortalamasının örnekleme dağılımı $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$ olup $z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$ olur.

Teorem (Merkezi Limit): μ ortalamalı σ^2 varyanslı bir kitleden $n \geq 30$ olacak şekildeki örneklemin ortalamasının örnekleme dağılımı $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$ sahiptir. O halde $z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$ olur.

Örnek: Elektronik parça üretimi yapan bir firma ortalaması 100 Ω ve standart sapması 10 Ω olan rezistanslar imal etmektedir. Rezistansın dağılımı normaldir. Örneklem hacmi 25 olan rezistansların oluşturduğu örneklemin ortalamasının 95 Ω değerinin altında olması ihtimali nedir?

Çözüm: Normal dağılıma sahip ortalaması ve varyansı bilinen kitleden alınan örneklem ortalamasının dağılımı da normal olup $n = 25$, $\mu_{\bar{X}} = 100$ ve $\sigma_{\bar{X}} = 10/\sqrt{25}$ 'tir.

$$P(\bar{X} < 95) = P\left(Z < \frac{95-100}{10/\sqrt{25}}\right) = P(Z < -2.5) = 1 - P(Z < 2.5) = 1 - 0.9938$$
$$= 0.0062.$$

Örnek: Demir bir testerenin bilendikten sonraki kullanım ömrü ortalaması 60 saat ve varyansı 36 olan dağılıma sahip olduğu biliniyor. İmal edilen bu testerelerden rasgele seçilen 64 birimlik örneklemin ortalamasının 58 ile 61 saat arasında olması olasılığı nedir?

Çözüm: Kitle normal olmasa da çekilen örneklem büyüklüğü $n = 64 \geq 30$ olup $\bar{X} \sim N(\mu = 60, \frac{\sigma^2}{n} = \frac{36}{64})$ olur. O halde bizden istenilen olasılık

$$P(58 < \bar{X} < 61) = P\left(\frac{58-60}{6/\sqrt{64}} < Z < \frac{61-60}{6/\sqrt{64}}\right) = P(-2.67 < Z < 1.33)$$
$$= P(Z < 1.33) - \big(1 - P(Z < 2.67)\big) = 0.9082 - (1 - 0.9962)$$
$$= 0.9044$$

Örnek: X sürekli rasgele değişkeninin olasılık yoğunluk fonksiyonu

$$f(x) = \begin{cases} 1/4 & 0 \leq x \leq 4 \\ 0 & d.d \end{cases}$$

şeklindedir. Bu kitleden rasgele seçilen 50 birimlik örneklemin ortalamasının 2.5 'tan az olması olasılığı nedir?

Çözüm: Örneklemin gözlem sayısı $n = 50 \geq 30$ olup örneklem ortalamasının örnekleme dağılımı normal dağılıma sahiptir. Öncelikle kitlenin ortalaması ve varyansını bulalım. X r.d. düzgün dağılıma sahip olup $\mu = \frac{0+4}{2} = 2$ ve $\sigma^2 = \frac{(4-0)^2}{12} = \frac{4}{3} = 1.33$'tür. Ya da

$$\mu = E(X) = \int_0^4 x * \frac{1}{4} dx = \frac{1}{4} * \frac{x^2}{2}\Big|_0^4 = 2$$

$$\sigma^2 = V(X) = E(X^2) - E(X)^2 = \int_0^4 x^2 * \frac{1}{4} dx - 2^2 = \frac{1}{4} * \frac{x^3}{3}\Big|_0^4 = \frac{16}{3} - 4 = \frac{4}{3} = 1.33$$

O halde $\mu_{\bar{X}} = 2$ ve $\sigma_{\bar{X}}^2 = \frac{1.33}{50}$ olup $\bar{X} \sim N(2, \frac{1.33}{50})$ olur.

$$P(\bar{X} < 2.5) = P\left(Z < \frac{2.5-2}{\sqrt{1.33}/\sqrt{50}}\right) = P(Z < 3.07) = 0.9989$$

10. PARAMETRE TAHMİNİ

Parametre tahminini nokta tahmini ve aralık tahmini olarak iki kısımda incelenir.

10.1. **Nokta Tahmini**: θ kitlenin bilinmeyen parametresi olmak üzere, örneklemden elde

edilen $\hat{\theta}$ değerine θ'nın tahmincisi denir.

Varsayalım ki bir mühendis otomobil şasisi üzerinde kullanılan bir parçanın kopma mukavemetini analiz ediyor olsun. Mühendis parçaların oluşturduğu kitleye ait ortalama kopma mukavemetini tahmin etmek için örneklem verisinin ortalamasını kullanacaktır. Bu sayı nokta tahmini olarak bilinir. Örneğin μ'nün nokta tahmini $\bar{x}$ olup $\hat{\mu} = \bar{x}$ şeklinde gösterilir.

Örnek: Bir bilgisayar şirketinin ürettiği ekran kartlarından rasgele seçilen 5 tanesinin ömürleri 12, 15, 18, 20 ve 24 ay olarak bulunuyor.

Bu bilgiye dayanarak şirketin ürettiği tüm ekran kartlarının ortalama ömrünün nokta tahmini $\hat{\mu} = \bar{x} = \frac{12+15+18+20+24}{5} = 17.80$ ay olarak hesaplanır.

Benzer şekilde kitlenin varyansı σ^2 nin nokta tahmini de örneklemin varyansı olup $\widehat{\sigma^2} = s^2 = 21.2$ 'dir.

10.2. **Aralık Tahmini (Güven Aralığı):** Kitle parametresi θ'yı nokta tahminleme yerine belirli bir hata payı (e) içeren bir aralıkla da tahmin edebiliriz. O halde $\left[\hat{\theta} - e,\ \hat{\theta} + e\right]$ aralığına θ nın aralık tahmini denir.

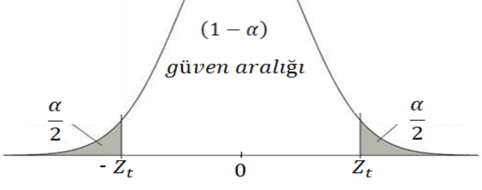
$$P\left[\hat{\theta} - e \le \theta \le \hat{\theta} + e\right] = 1 - \alpha$$

eşitliğinde $(1-\alpha)$ değeri, θ nın $\left[\hat{\theta} - e,\ \hat{\theta} + e\right]$ aralığında bulunması olasılığını gösterir. Burada α anlam seviyesi genellikle 0.01, 0.05 ve 0.10 alınır ve sırasıyla kitle parametresi için %99, %95 ve %90 güven aralıkları belirlenmiş olur.

Örneğin $\alpha = 0.05$ alınırsa bunun anlamı kitle parametresinin %95 güven aralığı demek olur ve 100 tekrarda elde edilen tahminin en az 95'i belirlenen aralıkta olduğunun göstergesidir.

10.2.1. **Kitle ortalaması ($\boldsymbol{\mu}$) için aralık tahmini:** Kitle ortalamasının aralık tahmini için durum söz konusudur. Bu durumlar kitlenin varyansının bilinip bilinmediği ve kitlenin normal dağılıma sahip olup olmadığıdır.

a) Kitlenin varyansı (σ^2) biliniyorsa ve kitle normal dağılıma sahipse örneklem büyüklüğü (n) ne olursa olsun ortalamanın örnekleme dağılımı da normal olup $Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$.



$$P(-z_t \le Z \le z_t) = 1 - \alpha$$

$$P\left(-z_t \le \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \le z_t\right) = 1 - \alpha$$

$$P\left(\bar{X} - z_t\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{X} + z_t\frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha.$$

Burada $z_t = z_{1-\frac{\alpha}{2}}$ dir. Çünkü z_t değerinden eşit ve küçük olması olasılıkları toplam

$(1-\alpha) + \frac{\alpha}{2} = 1 - \frac{\alpha}{2}$ dir.

Örneğin; $\alpha = 0.05$ için $z_t = z_{1-\frac{0.05}{2}} = z_{0.975} = 1.96$

O halde μ nün $1-\alpha$ güven arağı $\bar{x} - \left[z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right] \leq \mu \ \leq \bar{x} + \left[z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right]$ olarak bulunur.

Örnek: Varyansı 225 olan normal dağılmış bir kitleden 25 birim büyüklüğünde örneklem çekilmiş ve ortalaması 110 olarak bulunmuştur. Kitle ortalamasının %95 güven aralığını belirleyiniz.

Çözüm: $\sigma^2 = 225$ biliniyor ve kitle normal olup örneklemin gözlem sayısına bakılmaksızın kitle ortalamasının aralık tahmini $\bar{x} \mp \left[z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right]$ dur. $\sigma^2 = 225$ olup $\sigma = 15$, $\bar{x} = 110$ ve $n = 25$ dir. $1-\alpha = 0.95$ ise $\alpha = 0.05$ olup $z_{1-\frac{\alpha}{2}} = z_{1-\frac{0.05}{2}} = z_{0.975} = 1.96$ dır.

O halde μ nün %95 güven aralığı

$$110 - \left[1.96 * \frac{15}{\sqrt{25}}\right] \leq \mu \ \leq 110 + \left[1.96 * \frac{15}{\sqrt{25}}\right] \Rightarrow 104.12 \leq \mu \ \leq 115.88$$

Yani kitle ortalaması %95 ihtimalle 104.12 ile 115.88 arasındadır.

b) Kitlenin varyansı (σ^2) biliniyor ve kitlenin dağılımı normal değilse örneklem büyüklüğü $n \geq 30$ olduğunda merkezi limit teoreminden dolayı ortalamanın örnekleme dağılımı da normal olup $Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$ dir. Yine kitle ortalaması μ nün aralık tahmini $\bar{x} - \left[z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right] \leq \mu \ \leq \bar{x} + \left[z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right]$ dir.

Örnek: Önceki bilgilere dayanarak ehliyet sınavına giren adayların aldıkları puanların standart sapması 25 puan olduğu bilinmektedir. Bu sınava giren adaylardan rasgele seçilen 80 adayın puan ortalaması 75 olduğuna göre kitlenin ortalama puanının %90 güven aralığını belirleyiniz.

Çözüm: Standart sapma $\sigma = 25$ olup varyans biliniyor demektir. Lakin kitlenin dağılımı bilinmediğinden güven aralığı için sadece örneklemin gözlem sayısı $n \geq 30$ olması durumunda Z tablo değerleri kullanılabilir. Burada $n = 80 \geq 30$ olup güven aralığı

$\bar{x} \mp \left[z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right]$ ile belirlenebilir. $\bar{x} = 75$, $1-\alpha = 0.90$ olup $\alpha = 0.10$ dir.

Buradan $z_{1-\frac{0.10}{2}} = z_{0.95} = 1.64$ dır. O halde μ'nün %90 güven aralığı

$$75 - \left[1.64 * \frac{25}{\sqrt{80}}\right] \leq \mu \ \leq 75 + \left[1.64 * \frac{25}{\sqrt{80}}\right] \Rightarrow 70.42 \leq \mu \ \leq 79.58$$

Yani kitle puan ortalaması %90 ihtimalle 70.42 ile 79.58 arasındadır.

c) Kitlenin varyansı (σ^2) bilinmiyorsa ve kitle normal dağılıma sahipse kitlenin varyansı (σ^2) yerine onun nokta tahmini olan örneklem varyansı (s^2) kullanılır. μ'nün $1-\alpha$ güven aralığı örneklemdeki gözlem sayısına göre sayısına (n) göre;

$n \geq 30$ ise $Z = \frac{\bar{X}-\mu}{s/\sqrt{n}} \sim N(0,1)$ olup $\bar{x} - \left[z_{1-\frac{\alpha}{2}} \frac{s}{\sqrt{n}}\right] \leq \mu \ \leq \bar{x} + \left[z_{1-\frac{\alpha}{2}} \frac{s}{\sqrt{n}}\right]$

$n < 30$ ise $T = \frac{\bar{X}-\mu}{s/\sqrt{n}} \sim t_{n-1}$ olup $\bar{x} - \left[t_{1-\frac{\alpha}{2};\,(n-1)} \frac{s}{\sqrt{n}}\right] \leq \mu \ \leq \bar{x} + \left[t_{1-\frac{\alpha}{2};\,(n-1)} \frac{s}{\sqrt{n}}\right]$

ile belirlenir. Burada $t_{1-\frac{\alpha}{2};\,(n-1)}$ değeri $n-1$ serbestlik dereceli t dağılımının $1-\frac{\alpha}{2}$ olasılığına sahip tablo değeridir.

Örnek: Normal dağılan bir kitleden alınan 36 birimlik örneklemin ortalaması 25 ve varyansı 16 olduğuna göre kitle ortalamasının %95 güven aralığını belirleyiniz.

Çözüm: σ^2 bilinmiyor lakin ve kitle normal olup $n = 36 \geq 30$ olduğundan σ yerine onun tahmini s yi alarak kitle ortalamasının aralık tahmini $\bar{x} \mp \left[z_{1-\frac{\alpha}{2}} \frac{s}{\sqrt{n}}\right]$ ile hesaplanır. Burada $\bar{x} = 25$, $s^2 = 16$ ve $\alpha = 0.05$ olup $z_{1-\frac{\alpha}{2}} = z_{1-\frac{0.05}{2}} = z_{0.975} = 1.96$ dir.

$$25 - \left[1.96 * \frac{4}{\sqrt{36}}\right] \leq \mu \ \leq 25 + \left[1.96 * \frac{4}{\sqrt{36}}\right] \Rightarrow 23.69 \leq \mu \ \leq 26.31$$

Yani kitle ortalaması %95 ihtimalle 23.69 ile 26.31 arasındadır.

Örnek: Normal dağıldığı bilinen bir kitleden alınan 16 birimlik bir örneklemin ortalaması 35 ve standart sapması 4.6 olarak bulunmuştur. Kitle ortalamasının α=0.05 anlam seviyesinde güven aralığını belirleyiniz.

Çözüm: σ^2 bilinmiyor lakin ve kitle normal olup n=16<30 olduğundan σ yerine onun tahmini s'yi alarak kitle ortalamasının aralık tahmini $\bar{x} \mp \left[t_{1-\frac{\alpha}{2};\,(n-1)} \frac{s}{\sqrt{n}}\right]$ ile hesaplanır. Burada $\bar{x}$=35, s=4.6 ve α=0.05 dir. Bizden %95 güven aralığı istenmektedir$t_{1-\frac{\alpha}{2};\,(n-1)} =$
$t_{1-\frac{0,05}{2};\,(16-1)}$=2.1314 olup

$$35-[2.1314*\frac{4.6}{\sqrt{16}}]\leq\mu \leq 35+[2.1314*\frac{4.6}{\sqrt{16}}]\Rightarrow 32.55\leq\mu \leq 37.45$$

Yani kitle ortalaması %95 ihtimalle 32.55 ile 37.45 arasındadır.

Kitlenin varyansı (σ^2) bilinmiyorsa ve kitle normal değilse uygulama açısından sadece n≥30 için
$\bar{x} - \left[z_{1-\frac{\alpha}{2}} \frac{s}{\sqrt{n}}\right] \leq \mu \leq \bar{x} + \bar{x} + \left[z_{1-\frac{\alpha}{2}} \frac{s}{\sqrt{n}}\right]$ dir.

Örnek: Bir otelin lobisine gelen müşterilerin ortalama kaç dakika burada vakit geçirdikleri araştırmak istenmektedir. Rasgele seçilen 36 kişinin lobide kaç dakika kaldıkları gözlemlenmiş ve kayıt altına alınmıştır. Bu verilerin ortalaması 25 ve standart sapması 4 olarak hesaplanmıştır. Lobiye gelen tüm müşterilerin ortalama geçirdikleri sürenin %95 güven aralığını bulunuz.

Çözüm: σ^2 ve kitlenin dağılımı bilinmiyor ancak $n = 36 \geq 30$ olduğundan kitle ortalamasının aralık tahmini $\bar{x} \mp \left[z_{1-\frac{\alpha}{2}} \frac{s}{\sqrt{n}}\right]$ ile hesaplanır. Burada $\bar{x} = 25$, $s = 4$ ve

$1 - \alpha = 0.95$ olup $\alpha = 0.05$'tir. Buradan $z_{1-\frac{\alpha}{2}} = z_{1-\frac{0.05}{2}} = z_{0.975} = 1.96$ olup

$$25 - \left[1.96 * \frac{4}{\sqrt{36}}\right] \leq \mu \ \leq 25 + \left[1.96 * \frac{4}{\sqrt{36}}\right] \Rightarrow 23.69 \leq \mu \ \leq 26.31.$$

Yani %95 ihtimalle müşteriler lobide ortalama 23.69 ile 26.31 dakika arasında vakit geçirmektedir.

Not: μ nün $1 - \alpha$ güven aralığını şöyle özetleyebiliriz:

a) σ^2 biliniyor ve kitle normal ise n ne olursa olsun $\bar{x} \mp \left[z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right]$

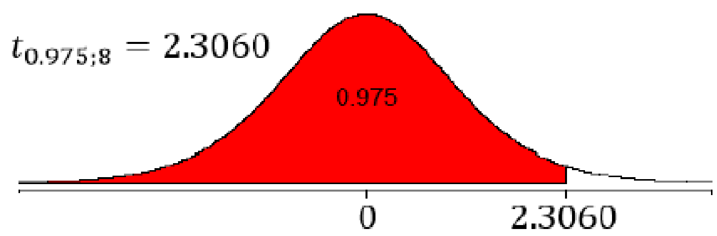
b) σ^2 biliniyor ve kitle normal değilse ancak $n \geq 30$ ise $\bar{x} \mp \left[z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right]$

c) σ^2 bilinmiyor ve kitle normal olsun ya da olmasın $n \geq 30$ ise $\bar{x} \mp \left[z_{1-\frac{\alpha}{2}} \frac{s}{\sqrt{n}}\right]$

d) σ^2 bilinmiyor ve kitle normal olsun $n < 30$ ise $\bar{x} \mp \left[t_{1-\frac{\alpha}{2};\,(n-1)} \frac{s}{\sqrt{n}}\right]$

Birikimli (Kümülatif) Student's t - Dağılım Tablosu

s.d.	$t_{0.70}$	$t_{0.80}$	$t_{0.90}$	$t_{0.95}$	$t_{0.975}$	$t_{0.99}$	$t_{0.995}$	$t_{0.999}$	$t_{0.9995}$
1	0.7265	1.3764	3.0777	6.3138	12.706	31.821	63.657	318.31	636.62
2	0.6172	1.0607	1.8856	2.9200	4.3027	6.9646	9.9248	22.327	31.599
3	0.5844	0.9785	1.6377	2.3534	3.1824	4.5407	5.8409	10.215	12.924
4	0.5686	0.9410	1.5332	2.1318	2.7764	3.7469	4.6041	7.1732	8.6103
5	0.5594	0.9195	1.4759	2.0150	2.5706	3.3649	4.0321	5.8934	6.8688
6	0.5534	0.9057	1.4398	1.9432	2.4469	3.1427	3.7074	5.2076	5.9588
7	0.5491	0.8960	1.4149	1.8946	2.3646	2.9980	3.4995	4.7853	5.4079
8	0.5459	0.8889	1.3968	1.8595	2.3060	2.8965	3.3554	4.5008	5.0413
9	0.5435	0.8834	1.3830	1.8331	2.2622	2.8214	3.2498	4.2968	4.7809
10	0.5415	0.8791	1.3722	1.8125	2.2281	2.7638	3.1693	4.1437	4.5869
11	0.5399	0.8755	1.3634	1.7959	2.2010	2.7181	3.1058	4.0247	4.4370
12	0.5386	0.8726	1.3562	1.7823	2.1788	2.6810	3.0545	3.9296	4.3178
13	0.5375	0.8702	1.3502	1.7709	2.1604	2.6503	3.0123	3.8520	4.2208
14	0.5366	0.8681	1.3450	1.7613	2.1448	2.6245	2.9768	3.7874	4.1405
15	0.5357	0.8662	1.3406	1.7531	2.1314	2.6025	2.9467	3.7328	4.0728
16	0.5350	0.8647	1.3368	1.7459	2.1199	2.5835	2.9208	3.6862	4.0150
17	0.5344	0.8633	1.3334	1.7396	2.1098	2.5669	2.8982	3.6458	3.9651
18	0.5338	0.8620	1.3304	1.7341	2.1009	2.5524	2.8784	3.6105	3.9216
19	0.5333	0.8610	1.3277	1.7291	2.0930	2.5395	2.8609	3.5794	3.8834
20	0.5329	0.8600	1.3253	1.7247	2.0860	2.5280	2.8453	3.5518	3.8495
21	0.5325	0.8591	1.3232	1.7207	2.0796	2.5176	2.8314	3.5272	3.8193
22	0.5321	0.8583	1.3212	1.7171	2.0739	2.5083	2.8188	3.5050	3.7921
23	0.5317	0.8575	1.3195	1.7139	2.0687	2.4999	2.8073	3.4850	3.7676
24	0.5314	0.8569	1.3178	1.7109	2.0639	2.4922	2.7969	3.4668	3.7454
25	0.5312	0.8562	1.3163	1.7081	2.0595	2.4851	2.7874	3.4502	3.7251
26	0.5309	0.8557	1.3150	1.7056	2.0555	2.4786	2.7787	3.4350	3.7066
27	0.5306	0.8551	1.3137	1.7033	2.0518	2.4727	2.7707	3.4210	3.6896
28	0.5304	0.8546	1.3125	1.7011	2.0484	2.4671	2.7633	3.4082	3.6739
29	0.5302	0.8542	1.3114	1.6991	2.0452	2.4620	2.7564	3.3962	3.6594
30	0.5300	0.8538	1.3104	1.6973	2.0423	2.4573	2.7500	3.3852	3.6460
35	0.5292	0.8520	1.3062	1.6896	2.0301	2.4377	2.7238	3.3400	3.5911
40	0.5286	0.8507	1.3031	1.6839	2.0211	2.4233	2.7045	3.3069	3.5510
45	0.5281	0.8497	1.3006	1.6794	2.0141	2.4121	2.6896	3.2815	3.5203
50	0.5278	0.8489	1.2987	1.6759	2.0086	2.4033	2.6778	3.2614	3.4960
60	0.5272	0.8477	1.2958	1.6706	2.0003	2.3901	2.6603	3.2317	3.4602
70	0.5268	0.8468	1.2938	1.6669	1.9944	2.3808	2.6479	3.2108	3.4350
80	0.5265	0.8461	1.2922	1.6641	1.9901	2.3739	2.6387	3.1953	3.4163
90	0.5263	0.8456	1.2910	1.6620	1.9867	2.3685	2.6316	3.1833	3.4019
100	0.5261	0.8452	1.2901	1.6602	1.9840	2.3642	2.6259	3.1737	3.3905
200	0.5252	0.8434	1.2858	1.6525	1.9719	2.3451	2.6006	3.1315	3.3398
∞	0.5244	0.8416	1.2816	1.6449	1.9600	2.3263	2.5758	3.0902	3.2905



Örnek: Normal dağıldığı bilinen bir kitleden alınan örneklemin $\{1,2,3,4,5,6,7,8,9\}$ ölçüm değerlerini kullanarak kitle ortalamasının %95 güven aralığını bulunuz.

Çözüm: σ^2 bilinmiyor lakin ve kitle normal olup $n = 9 < 30$ olduğu için güven aralığı $\bar{x} \mp \left[t_{1-\frac{\alpha}{2};\,(n-1)} \frac{s}{\sqrt{n}}\right]$ ile hesaplanır. Burada örneklemin ortalaması ve standart sapmasının hesaplanması gerekmektedir.

$$\bar{x} = \frac{1+2+\cdots+9}{9} = 5 \;\Rightarrow s^2 = \frac{(1-5)^2+(2-5)^2+\cdots+(9-5)^2}{9-1} = 7.5 \;\Rightarrow s = 2.74$$

$$t_{1-\frac{\alpha}{2};\,(n-1)} = t_{1-\frac{0.05}{2};\,(9-1)} = t_{0.975;\,8} = 2.3060$$

$$5 - \left[2.3060 * \frac{2.74}{\sqrt{9}}\right] \le \mu \;\le 5 + \left[2.3060 * \frac{2.74}{\sqrt{9}}\right] \Rightarrow 2.89 \le \mu \;\le 7.11$$

olur. Kitle ortalaması %95 ihtimalle 2.89 ile 7.11 arasındadır.

Örnek: Bir araba üreticisi rasgele seçtiği 6 arabanın CO_2 emisyon değerlerini $\{100.1, 99.9, 100.7, 99.3, 99{,}5, 100.5\}$ g/km olarak ölçmüştür. Verilerin normal dağıldığı varsayımı altında üretilen arabaların ortalama CO_2 emisyon değerlerinin %90 güven aralığını bulunuz.

Çözüm: σ^2 bilinmiyor lakin ve kitle normal olup $n = 6 < 30$ olduğu için güven aralığı $\bar{x} \mp \left[t_{1-\frac{\alpha}{2};\,(n-1)} \frac{s}{\sqrt{n}}\right]$ ile hesaplanır.

$$\bar{x} = \frac{100.1+\cdots+100.5}{6} = 100$$

$$s^2 = \frac{(100.1-100)^2+\cdots+(100.5-100)^2}{6-1} = 0.3 \Rightarrow s = 0.55$$

$$t_{1-\frac{\alpha}{2};\,(n-1)} = t_{1-\frac{0.10}{2};\,(6-1)} = t_{0.95;\,5} = 2.0150$$

$$100 - \left[2.0150 * \frac{0.55}{\sqrt{6}}\right] \le \mu \;\le 100 + \left[2.0150 * \frac{0.55}{\sqrt{6}}\right] \Rightarrow 99.54 \le \mu \;\le 100.46$$

Yani üretilen arabaların ortalama CO_2 emisyon değerlerinin %90 güvenle 99.54 ile 100.46 arasındadır.

Örnek: Bir hastalığın tedavisinde yeni bir yöntem geliştirilmiştir. Rasgele seçilen 12 hasta sözü edilen bu yöntem ile tedavi edilerek iyileşinceye kadar geçen süreler $\{10, 12, 8, 9, 9, 17, 9, 14, 3, 7, 12, 10\}$ gün şeklindedir. Verilerin normal dağıldığı varsayımı altında $\alpha = 0.05$ anlam seviyesinde hastaların ortalama iyileşme süresi için güven aralığını bulunuz.

Çözüm: Kitle normal olup $n = 12 < 30$ olduğu t tablo değeri kullanılır.

$$\bar{x} = \frac{10+12+\cdots+10}{12} = 10 \Rightarrow s^2 = 12.55 \;\Rightarrow s = 3.54$$

$$t_{1-\frac{\alpha}{2};\,(n-1)} = t_{1-\frac{0.05}{2};\,(12-1)} = t_{0.975;\,11} = 2.2010$$

$$10-\left[2.2010*\frac{3.54}{\sqrt{12}}\right]\leq\mu\ \leq 10+\left[2.2010*\frac{3.54}{\sqrt{12}}\right]\Rightarrow 7.75\leq\mu\ \leq 12.24$$

O halde hastaların bu yöntem ile ortalama iyileşme süresi %95 güvenle 7.75 ile 12.24 gün arasındadır.

Örnek: Belirli bir mahalleden rasgele seçilen 10 binanın yaşlarına ilişkin verilen $\{3,3,3,4,4,5,6,7,7,8\}$ yıl olarak belirlenmiştir. Mahalledeki binaların yaş değerlerinin normal dağıldığı varsayımı altında mahalledeki binaların ortalama yaşlarına ilişkin $\alpha=0.01$ anlam seviyesinde güven aralığını bulunuz.

Çözüm: Kitle normal olup $n=10<30$ olduğu t tablo değeri kullanılır.

$$\bar{x}=\frac{3+3+\cdots+8}{10}=5\Rightarrow s^2=3.56\ \Rightarrow s=1.89$$

$$t_{1-\frac{\alpha}{2};\,(n-1)}=t_{1-\frac{0.01}{2};\,(10-1)}=t_{0.995;\,9}=3.2498$$

$$5-\left[3.2498*\frac{1.89}{\sqrt{10}}\right]\leq\mu\ \leq 5+\left[3.2498*\frac{1.89}{\sqrt{10}}\right]\Rightarrow 3.06\leq\mu\ \leq 6.94.$$

Örnek: Bir çay fabrikası ürettiği 1 kilogramlık çay paketlerinin ortalama ağırlığı için %95 güven aralığını belirlemek istemektedir. Üretilen çay paketlerinden rasgele seçilen 49 tanesinin ortalama ağırlığı 995 gr ve standart sapması 70 gr olduğuna göre güven aralığını oluşturunuz.

Çözüm: σ^2 ve kitlenin dağılımı bilinmiyor lakin $n=49\geq 30$ olduğundan kitle ortalamasının aralık tahmini $\bar{x}\mp\left[z_{1-\frac{\alpha}{2}}\frac{s}{\sqrt{n}}\right]$ ile hesaplanır.

Burada $\bar{x}=995$, $s=70$ ve $\alpha=0.05$ olup $z_{1-\frac{\alpha}{2}}=z_{0.975}=1.96$ olur.

$$995-\left[1.96*\frac{70}{\sqrt{49}}\right]\leq\mu\ \leq 995+\left[1.96*\frac{70}{\sqrt{49}}\right]\Rightarrow 975.40\leq\mu\ \leq 1014.60.$$

10.2.2. **Kitle oranı (*Π*) için aralık tahmini:** Bir fabrikada üretilen ürünlerin kusurlu ya da kusursuz olması, bir fakültedeki öğrencilerin başarılı ya da başarısız olması gibi iki sonuçlu kitlelerin oran parametresi $\Pi=R/N$ ile hesaplanır ($0\leq\Pi\leq 1$). Burada N kitledeki eleman sayısı olup R de ilgilenilen özelliğe sahip gözlem sayısıdır.

Kitleden alınan n boyutlu örneklemde ilgilenilen özelliğe sahip r tane gözlem varsa örneklemin oranı $p=r/n$ olacaktır. Eğer Π, 0 veya 1'e çok yakın değil ve n de yeterince büyükse ($np>5$ veya $n(1-p)>5$) örnek oranının dağılımı ortalaması $\mu=\Pi$ ve varyansı $\sigma_p^2=\Pi(1-\Pi)/n$ olan normal dağılıma sahiptir. Kitle oranı Π bilinmediği durumlarda tahmini olan p alınarak örnek varyansı $s_p^2=\frac{p(1-p)}{n}\Rightarrow s_p=\sqrt{\frac{p(1-p)}{n}}$ olur. O halde kitle oranı Π için $1-\alpha$ güven aralığı

$$p-\left[z_{1-\frac{\alpha}{2}}\,s_p\right]\leq\Pi\leq p+\left[z_{1-\frac{\alpha}{2}}\,s_p\right]$$

ile hesaplanır.

Örnek: Bir fabrikada üretilen ürünlerin içerisinden rasgele seçilen 80 ürünün 6 tanesinin bozuk olduğu gözlemlenmiştir. Fabrikada üretilen ürünlerin bozuk oranı için %95 güven aralığını bulunuz.

Çözüm: $n = 80,\ r = 6$ olup $p = \frac{r}{n} = \frac{6}{80} = 0.075$ ve $s_p = \sqrt{\frac{0.075(1-0.075)}{80}} = 0.0294$ dur. Burada $np = 80 * 0.075 = 6 > 5$ olup Π nin güven aralığı $p \mp \left[z_{1-\frac{\alpha}{2}} s_p\right]$ ile hesaplanabilir. $1-\alpha = 0.95$ ise $\alpha = 0.05$ dir. Buradan $z_{1-\frac{\alpha}{2}} = z_{1-\frac{0.05}{2}} = z_{0.975} = 1.96$ olup

$$0.075 - [1.96 * 0.0294] \leq \Pi \leq 0.075 + [1.96 * 0.0294] \Rightarrow 0.0174 \leq \Pi \leq 0.1326.$$

Yani üretilen ürünlerin bozuk olma oranı %95 güvenle (ihtimalle) %1.74 ile %13.26 arasındadır.

Örnek: Seçim öncesi yapılan bir ankette 1000 kişiden 200 ünün bir siyasi partiyi desteklediklerini tespit edilmiştir. Bu siyasi partinin oy oranı için %99 güven aralığını bulunuz.

Çözüm: $n = 1000,\ r = 200$ olup $p = \frac{r}{n} = \frac{200}{1000} = 0.20$ ve $s_p = \sqrt{\frac{0.20(1-0.20)}{1000}} = 0.0126$ dur. Burada $np = 200 > 5$ olup Π nin güven aralığı $p \mp \left[z_{1-\frac{\alpha}{2}} s_p\right]$ ile hesaplanabilir. $1-\alpha = 0.99$ olup $\alpha = 0.01$ dir. Buradan $z_{1-\frac{\alpha}{2}} = z_{1-\frac{0.01}{2}} = Z_{0.995} = 2.58$ bulunur. Son olarak

$$0.20 - [2.58 * 0.0126] \leq \Pi \leq 0.20 + [2.58 * 0.0126] \Rightarrow 0.1675 \leq \Pi \leq 0.2325$$

olup bu da siyasi partinin oy oranı %99 ihtimalle %17 ile %23 arasında olacağını söyler.

Örnek: Bir bölgede sigara içen insanlar arasında kanser hastası olanların oranının belirlenmesi istenmektedir. Bu kişiler arasından rasgele seçilen 1500 kişinin 375'i kanser hastası olduğu tespit edilmiştir. Bu bölgedeki sigara içen insanların kansere yakalanması oranı için %95 güven aralığını bulunuz.

Çözüm: $n = 1500,\ r = 375$ olup $p = \frac{375}{1500} = 0.25$ ve $s_p = \sqrt{\frac{0.25(1-0.25)}{1500}} = 0.0112$ dur. $\alpha = 0.05$ olup $z_{1-\frac{\alpha}{2}} = z_{1-\frac{0.05}{2}} = z_{0.975} = 1.96$ dır. Π'nin güven aralığı

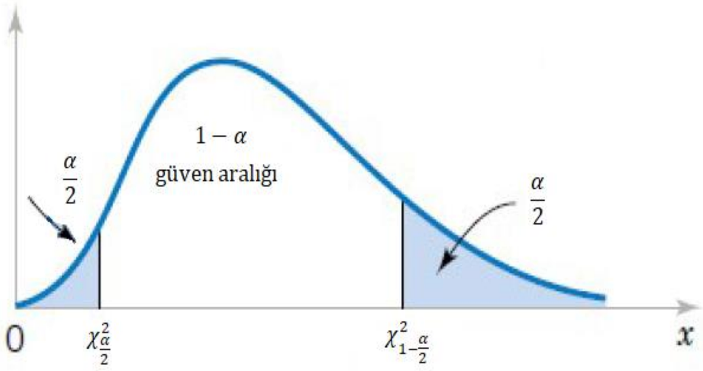
$$0.25 - [1.96 * 0.0112] \leq \Pi \leq 0.25 + [1.96 * 0.0112] \Rightarrow 0.2280 \leq \Pi \leq 0.2720.$$

O halde bu bölgedeki sigara içen insanların kansere yakalanması oranı %95 ihtimalle %23 ile %27 arasındadır.

10.2.3. **Kitle varyansı (σ^2) için aralık tahmini:** μ ortalamalı σ^2 varyanslı normal dağılıma sahip bir kitleden çekilen n büyüklüğündeki örneklem varyansı s^2 olmak üzere

$$X^2 = \frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{(n-1)}$$

değeri $(n-1)$ serbestlik dereceli χ^2 (Ki-Kare) dağılımına sahiptir. Burada kitle varyansı σ^2 için $1-\alpha$ güven aralığı



$$\frac{(n-1)s^2}{\chi^2_{1-\frac{\alpha}{2};\,(n-1)}} \le \sigma^2 \le \frac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2};\,(n-1)}}$$

ile hesaplanır.

Örnek: Normal dağılan bir kitleden çekilen 12 birimlik örneklemin ortalaması 20 ve varyansı 10 olarak hesaplanmıştır. Kitle varyansı için %90 güven aralığını belirleyiniz.

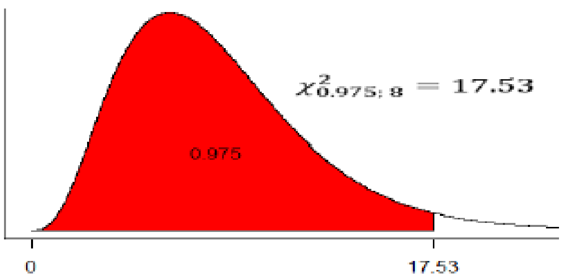
Çözüm: $s^2 = 10, n = 12, \alpha = 0.10$ olup $\chi^2_{\frac{\alpha}{2};\,(n-1)} = \chi^2_{\frac{0.10}{2};\,(12-1)} = \chi^2_{0.05;\,11} = 4.57$ ve $\chi^2_{1-\frac{\alpha}{2};\,(n-1)} = \chi^2_{1-\frac{0.10}{2};\,(12-1)} = \chi^2_{0.95;\,11} = 19.68$ olup σ^2 için %90 güven aralığı

$$\frac{(12-1)10}{19.68} \le \sigma^2 \le \frac{(12-1)10}{4.57} \Rightarrow 5.59 \le \sigma^2 \le 24.07$$

O halde %90 ihtimalle kitle varyansı σ^2 değeri 5.59 ile 24.07 arasındadır.

Birikimli (Kümülatif) Ki Kare Dağılım Tablosu

s.d	$\chi^2_{0.005}$	$\chi^2_{0.01}$	$\chi^2_{0.025}$	$\chi^2_{0.05}$	$\chi^2_{0.10}$	$\chi^2_{0.25}$	$\chi^2_{0.50}$	$\chi^2_{0.75}$	$\chi^2_{0.90}$	$\chi^2_{0.95}$	$\chi^2_{0.975}$	$\chi^2_{0.99}$	$\chi^2_{0.995}$
1	0.00	0.00	0.00	0.00	0.02	0.10	0.45	1.32	2.71	3.84	5.02	6.63	7.88
2	0.01	0.02	0.05	0.10	0.21	0.58	1.39	2.77	4.61	5.99	7.38	9.21	10.60
3	0.07	0.11	0.22	0.35	0.58	1.21	2.37	4.11	6.25	7.81	9.35	11.34	12.84
4	0.21	0.30	0.48	0.71	1.06	1.92	3.36	5.39	7.78	9.49	11.14	13.28	14.86
5	0.41	0.55	0.83	1.15	1.61	2.67	4.35	6.63	9.24	11.07	12.83	15.09	16.75
6	0.68	0.87	1.24	1.64	2.20	3.45	5.35	7.84	10.64	12.59	14.45	16.81	18.55
7	0.99	1.24	1.69	2.17	2.83	4.25	6.35	9.04	12.02	14.07	16.01	18.48	20.28
8	1.34	1.65	2.18	2.73	3.49	5.07	7.34	10.22	13.36	15.51	17.53	20.09	21.95
9	1.73	2.09	2.70	3.33	4.17	5.90	8.34	11.39	14.68	16.92	19.02	21.67	23.59
10	2.16	2.56	3.25	3.94	4.87	6.74	9.34	12.55	15.99	18.31	20.48	23.21	25.19
11	2.60	3.05	3.82	4.57	5.58	7.58	10.34	13.70	17.28	19.68	21.92	24.72	26.76
12	3.07	3.57	4.40	5.23	6.30	8.44	11.34	14.85	18.55	21.03	23.34	26.22	28.30
13	3.57	4.11	5.01	5.89	7.04	9.30	12.34	15.98	19.81	22.36	24.74	27.69	29.82
14	4.07	4.66	5.63	6.57	7.79	10.17	13.34	17.12	21.06	23.68	26.12	29.14	31.32
15	4.60	5.23	6.26	7.26	8.55	11.04	14.34	18.25	22.31	25.00	27.49	30.58	32.80
16	5.14	5.81	6.91	7.96	9.31	11.91	15.34	19.37	23.54	26.30	28.85	32.00	34.27
17	5.70	6.41	7.56	8.67	10.09	12.79	16.34	20.49	24.77	27.59	30.19	33.41	35.72
18	6.26	7.01	8.23	9.39	10.86	13.68	17.34	21.60	25.99	28.87	31.53	34.81	37.16
19	6.84	7.63	8.91	10.12	11.65	14.56	18.34	22.72	27.20	30.14	32.85	36.19	38.58
20	7.43	8.26	9.59	10.85	12.44	15.45	19.34	23.83	28.41	31.41	34.17	37.57	40.00
21	8.03	8.90	10.28	11.59	13.24	16.34	20.34	24.93	29.62	32.67	35.48	38.93	41.40
22	8.64	9.54	10.98	12.34	14.04	17.24	21.34	26.04	30.81	33.92	36.78	40.29	42.80
23	9.26	10.20	11.69	13.09	14.85	18.14	22.34	27.14	32.01	35.17	38.08	41.64	44.18
24	9.89	10.86	12.40	13.85	15.66	19.04	23.34	28.24	33.20	36.42	39.36	42.98	45.56
25	10.52	11.52	13.12	14.61	16.47	19.94	24.34	29.34	34.38	37.65	40.65	44.31	46.93
26	11.16	12.20	13.84	15.38	17.29	20.84	25.34	30.43	35.56	38.89	41.92	45.64	48.29
27	11.81	12.88	14.57	16.15	18.11	21.75	26.34	31.53	36.74	40.11	43.19	46.96	49.64
28	12.46	13.56	15.31	16.93	18.94	22.66	27.34	32.62	37.92	41.34	44.46	48.28	50.99
29	13.12	14.26	16.05	17.71	19.77	23.57	28.34	33.71	39.09	42.56	45.72	49.59	52.34
30	13.79	14.95	16.79	18.49	20.60	24.48	29.34	34.80	40.26	43.77	46.98	50.89	53.67
40	20.71	22.16	24.43	26.51	29.05	33.66	39.34	45.62	51.81	55.76	59.34	63.69	66.77
50	27.99	29.71	32.36	34.76	37.69	42.94	49.33	56.33	63.17	67.50	71.42	76.15	79.49
60	35.53	37.48	40.48	43.19	46.46	52.29	59.33	66.98	74.40	79.08	83.30	88.38	91.95
70	43.28	45.44	48.76	51.74	55.33	61.70	69.33	77.58	85.53	90.53	95.02	100.43	104.21
80	51.17	53.54	57.15	60.39	64.28	71.14	79.33	88.13	96.58	101.88	106.63	112.33	116.32
90	59.20	61.75	65.65	69.13	73.29	80.62	89.33	98.65	107.57	113.15	118.14	124.12	128.30
100	67.33	70.06	74.22	77.93	82.36	90.13	99.33	109.14	118.50	124.34	129.56	135.81	140.17



10.3. **Örneklem büyüklüğünün ($\boldsymbol{n}$) belirlenmesi:** Örneklem büyüklüğü arttıkça kitle parametrelerine ilişkin bilgilerimizin de artacağı kesindir. Fakat bu durumun maliyet ve zaman açısından dezavantaja sahiptir. O halde optimum örneklem büyüklüğünün belirlenmesi önemlidir. Kitle parametresi için belirlenen maksimum hata değeri e olmak üzere $\left[\hat{\theta} - e \le \theta \le \hat{\theta} + e\right]$ güven aralığına düşen kitle parametresi θ için örneklem sayısını belirleyelim.

10.3.1. **Kitle ortalamasının tahmini için örneklem büyüklüğü:** Kitle parametresi μ için e hata değerine sahip güven aralığı

$$\bar{x} - e \le \mu \le \bar{x} + e$$

olarak yazılabilir. Kitle varyansının bilindiği ve kitlenin normal dağıldığı durumda μ'nün $1-\alpha$ güven aralığı

$$\bar{x} - \left[z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right] \le \mu \le \bar{x} + \left[z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right]$$

olup burada hata değerini $e = \left[z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right]$ alınabilir. Bu eşitlikten n değeri çekilirse $n = \left[z_{1-\frac{\alpha}{2}}\frac{\sigma}{e}\right]^2$ bulunur. O halde μ nün e hata payına sahip $1-\alpha$ güven aralığında minimum örneklem büyüklüğü $n = \left[z_{1-\frac{\alpha}{2}}\frac{\sigma}{e}\right]^2$ alınabilir. Ancak burada kitle varyansının bilindiği bilgisine ihtiyaç vardır. Bu bilgi önceki çalışmalar yardımıyla temin edilebilir ya da σ yerine bir ön örnekleme yapılıp örneklemin standart sapması s kullanılabilir.

Örnek: Bir fabrikada üretilen çelik kabloların ortalama kaç tonluk bir gerilime dayanabilecekleri araştırılmaktadır. Yöneticiler, kabloların dayanma güçlerinin standart sapması 1 ton olduğunu ve normal dağıldıklarını geçmiş verilerden bilmektedir. Örneklemden hesaplanacak ortalama değer ile gerçek ortalama arasındaki farkın en fazla 0.5 ton olmasını istediklerine göre %90 güven aralığı oluşturmak için kaç kablonun test edilmesi gerekmektedir?

Çözüm: $\alpha = 0.10 \Rightarrow z_{1-\frac{\alpha}{2}} = Z_{0.95} = 1.64$, $e = |\mu - \bar{x}| = 0.5$, $\sigma = 1$ olup

$$n = \left[z_{1-\frac{\alpha}{2}}\frac{\sigma}{e}\right]^2 = \left[1.64 * \frac{1}{0.5}\right]^2 = 10.76 \cong 11.$$

10.3.2. **Kitle oranının tahmini için örneklem büyüklüğü:** Kitle parametresi Π nin $1-\alpha$ güven aralığı formülünde e hata değeri olmak üzere $e = \left[z_{1-\frac{\alpha}{2}} s_p\right] = \left[z_{1-\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{n}}\right]$ alınabilir. Bu eşitlikten n değeri çekilirse $n = \left[z_{1-\frac{\alpha}{2}}\frac{\sqrt{p(1-p)}}{e}\right]^2$ bulunur. O halde Π nin maksimum e hata payına sahip $1-\alpha$ güven aralığında $n = p(1-p)\left[\frac{z_{1-\frac{\alpha}{2}}}{e}\right]^2$ örneklem büyüklüğü olarak alınabilir. Ancak burada örneklem oranı p bilinmediği için geçmiş verilere ilişkin bilgiler ya da bir ön örnekleme yapılıp p tahmin edilebilir. Herhangi bir bilgiye sahip değilsek $p = 1/2$

alınırsa $p(1-p) = 1/4$ değeri maksimum olacağından n nin maksimum değerini bulmuş oluruz. Böylece $n = \frac{1}{4}\left[\frac{z_{1-\frac{\alpha}{2}}}{e}\right]^2$ alınabilir.

Örnek: Bir fabrikada üretilen ampullerin defolu olma yüzdesi yüzdesi tahmin edilmek isteniyor. En fazla %3 lük bir hata payı ile %95 güven aralığında

- Defolu olma oranının geçmiş bilgilere göre yaklaşık %10 bilgisiyle
- Defolu olma oranına ilişkin herhangi bir bilginin olmadığında durumlar için örneklem büyüklüğünü belirleyiniz.

Çözüm: $\alpha = 0.05$ olup $z_{1-\frac{\alpha}{2}} = 1.96$ ve $e = 0.03$ dür.

- Geçmiş bilgilere göre $p = 0.10$ kullanılırsa örneklem büyüklüğü

$$n = p(1-p)\left[\frac{z_{1-\frac{\alpha}{2}}}{e}\right]^2 = 0.10(1-0.10)\left[\frac{1.96}{0.03}\right]^2 = 384.16 \cong 384.$$

- Herhangi bir bilgiye sahip değilsek örneklem büyüklüğü

$$n = \frac{1}{4}\left[\frac{z_{1-\frac{\alpha}{2}}}{e}\right]^2 = \frac{1}{4}\left[\frac{1.96}{0.03}\right]^2 = 1067.11 \cong 1067.$$