BÖLÜM 11

TAHMİN

Örneklem: Birbirinden bağımsız ve aynı dağılıma sahip n tane $X_1, X_2, ..., X_n$ rastgele değişkenlerine bu dağılımdan alınan ''örneklem'' denir. n sayısına ''örneklem hacmi'' denir. Ölçme sonucunda bu rastgele değişkenlerin aldığı değerlere 'gözlem değerleri' ya da 'veri' denir.

İstatistik: Örneklemin bilinmeyen parametre içermeyen fonksiyonuna denir. Her istatistik bir rastgele değişkendir.

Tahmin Ediciler: Bir istatistik bilinmeyen bir parametrenin belirlenmesi amacıyla kullandığında bu istatistiğe ''tahmin edici'' denir.

Tahmin: Bilinmeyen bir parametre için önerilen tahmin edicinin aldığı değere ''tahmin'' denir.
$\hat{\theta}$ ile gösterilir.

Nokta Tahmini: Parametreye uygun istatistiğin tek sayısal değeridir. Örneğin; kitle ortalamasını tahmin etmek için örneklem ortalaması alınırsa, parametrenin nokta tahmin edicisi kullanılıyor demektir.

Aralık Tahmini: Bilinmeyen kitle parametresinin belli bir olasılıkla içinde bulunacağı rastgele iki sınırı belirlemektedir. Buna ''güven aralığı'' denir.

Merkezi Limit Teoremi: Ortalaması μ, varyansı σ^2 olan, herhangi bir kitleden yerine koyarak n birimlik örneklemin ortalamasına ait örneklem dağılımı ortalaması (μ), varyansı (σ^2 / n) olan normal dağılıma yaklaşır.

$X \sim N\left(\mu, \sigma^2\right)$ ise $\bar{X} \sim N(\mu, \sigma^2 / n)$ dir ve

$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \sim N(0,1)$$

şeklinde standart normal dağılıma dönüşür.

Güven Aralığı: Bilinmeyen θ kitle parametresini içine alan bir aralık, belirli bir olasılıkla tahmin edilebilir. Bu tahmine aralık tahmini, tahmin edilen aralığa da ''güven aralığı'' denir.

$$P(\hat{\theta}_1 < \theta < \hat{\theta}_2) = 1 - \alpha \ , \ \hat{\theta}_1 < \hat{\theta}_2 \qquad (\textit{iki yönlü güven aralığı})$$

Şeklinde gösterilir.

Tek taraflı güven aralığı

$$P(\theta < \hat{\theta}_2) = 1 - \alpha \ \ \textit{ya da} \ \ P(\theta > \hat{\theta}_1) = 1 - \alpha$$

şeklindedir.

Hipotez Testleri ve Güven Aralıkları

Geçerliliği olasılık esaslarına göre araştırılabilen ve bir karar verebilmek için öne sürülen varsayımlara ‘’hipotez’’ denir. Hipotez, kitle dağılımı ile ilgili öne sürülen bir önermedir. Örneklem dağılımlarından elde edilen istatistiklere bağlı kalarak örneklem dağılımının parametresi bilinen kitleye ait olup olmadığı araştırılır.
Hipotezlerin örneklem yardımıyla incelenmesine ‘’hipotez testi’’ denir. Test etmek istenilen hipotez H_0 ile karşıt hipotez ise H_1 ile gösterilir. Bir hipotez testi yapılıp H_0 hipotezinin kabulüne ya da reddine karar verildiğinde aşağıdaki dört durumdan birisi söz konusu olacaktır.

H_0 : *Yokluk hipotezi*

H_1 : *Alternatif hipotez*

Hipotez Testinde Kararlar:

	H_0 *doğru*	H_0 *yanlış*
H_0 *red edilemez*	*Uygun karar*	*II.tip hata*
H_0 *red*	*I.tip hata*	*Uygun karar*

I.Tip Hata: H_0 hipotezi gerçekte doğruyken H_0 hipotezi red edilirse I. Tip hata oluşur. Bu hatanın oluşması olasılığı α ’dır. Yani,

$$P(I.tip\ hata) = P(H_0\ red\ /\ H_0\ doğru) = \alpha$$

Testin anlamlılık düzeyi olarak anlamlandırılır. Çoğunlukla $\alpha = 0.05,\ 0.01, \ldots$ gibi seçilir.

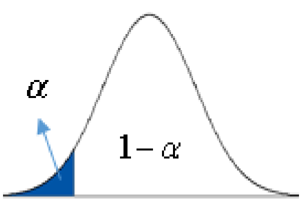
II.Tip Hata: H_0 hipotezi gerçekte yanlış iken H_0 hipotezi red edilmezse II. tip hata oluşur. Bu hatanın oluşması olasılığı β ’dır.

$$P(II.tip\ hata) = P(H_0\ red\ edilemez\ /\ H_0\ yanlış) = \beta$$

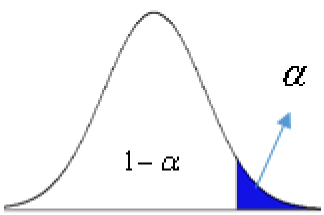
Testin Gücü: H_0 hipotezi gerçekte yanlış iken H_0 hipotezi red etme olasılığına testin gücü denir. Yani, testin gücü $P(H_0\ red\ /\ H_0\ yanlış) = 1-\beta$ ’dir.

*Hipotezin red edildiği bölgeye “red bölgesi” ya da “ kritik bölge’’ denir. (Taralı bölgeler kurulan hipoteze göre red bölgelerini göstermektedir.)

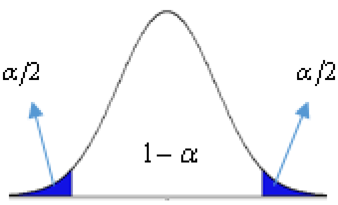
$H_0 : \mu = \mu_0$
$H_1 : \mu < \mu_0$



$H_0 : \mu = \mu_0$
$H_1 : \mu > \mu_0$



$H_0 : \mu = \mu_0$
$H_1 : \mu \neq \mu_0$



<u>Kitle Ortalaması için Hipotez Testi ve Güven Aralığı</u>

1) <u>Kitle varyansı σ^2 biliniyor</u>

Kitle varyansı σ^2 bilindiğinde kitle ortalaması için hipotez testi.

$$X_1, X_2, \dots, X_n \sim N(\mu, \sigma^2)$$

$$\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$$

$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$$

1) Hipotez kurulur.

$H_0 : \mu = \mu_0$	$H_0 : \mu = \mu_0$	$H_0 : \mu = \mu_0$
$H_1 : \mu < \mu_0$	$H_1 : \mu > \mu_0$	$H_1 : \mu \neq \mu_0$

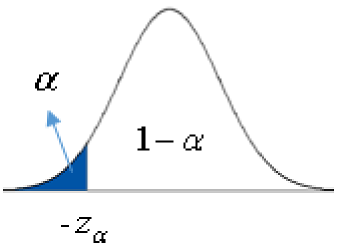
2) Test istatistiği hesaplanır.

$$Z_t = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$$

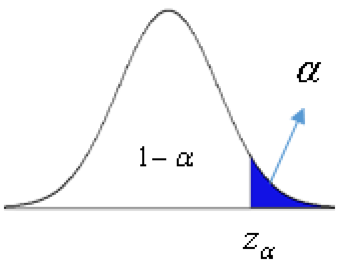
3) Karar verilir.

α anlam düzeyine göre kritik bölge belirlenerek hipotezin red edilip edilmeyeceğine karar verilir.

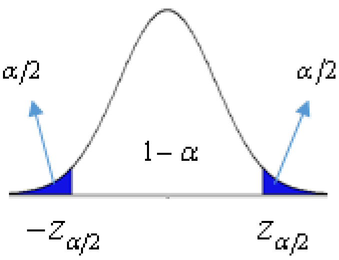
$H_0 : \mu = \mu_0$	$H_0 : \mu = \mu_0$	$H_0 : \mu = \mu_0$
$H_1 : \mu < \mu_0$	$H_1 : \mu > \mu_0$	$H_1 : \mu \neq \mu_0$



$z_t < -z_\alpha$ olduğunda H_0 hipotezi red edilir



$z_t > z_\alpha$ olduğunda H_0 hipotezi red edilir



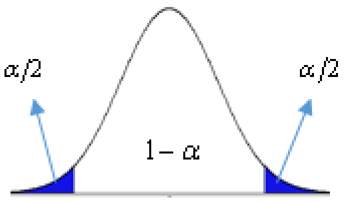
$z_t < -z_{\alpha/2}$ olduğunda

ya da

$z_t > z_{\alpha/2}$ olduğunda

H_0 hipotezi red edilir

Güven Aralığı



$$P(-z_{\alpha/2} < Z < z_{\alpha/2}) = 1-\alpha$$

$$P(-z_{\alpha/2} < \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} < z_{\alpha/2}) = 1-\alpha$$

$$P(\bar{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}) = 1-\alpha$$

$$\bar{x} \mp z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

p değeri: H_0 hipotezi doğru olduğunda, test istatistiğinin hesaplanan değerine eşit ya da daha uç değerler alması olasılığıdır. Bu olasılığa ‘’p değeri’’ denir.

$p < \alpha$ *ise* H_0 *hipotezi red edilir.*

$$p = \begin{cases} 2[1-P(Z \le Z_t)], & H_1 : \mu \ne \mu_0 \\ 1-P(Z \le Z_t), & H_1 : \mu > \mu_0 \\ P(Z \le Z_t), & H_1 : \mu < \mu_0 \end{cases}$$

Örnek: $100\,cm^3$ idrardaki keratin miktarı ölçülmüş ve ortalaması $\mu = 1.58\ mg$, varyansı $\sigma^2 = 0.05\ mg$ olduğu saptanmıştır. Fakat bir araştırmacı ortalama keratin miktarının daha fazla olduğunu iddia etmektedir. Bu amaçla rastgele 84 kişi seçilmiştir. Veriler aşağıdaki gibidir;

1.22	1.22	1.23	1.25	1.26	1.28	1.31	1.33	1.33	1.36
1.37	1.38	1.40	1.40	1.40	1.43	1.43	1.43	1.44	1.46
1.46	1.46	1.47	1.47	1.47	1.47	1.49	1.49	1.50	1.51
1.51	1.52	1.52	1.52	1.53	1.54	1.54	1.55	1.56	1.56
1.57	1.57	1.58	1.58	1.58	1.59	1.59	1.60	1.60	1.61
1.62	1.65	1.65	1.66	1.66	1.66	1.66	1.67	1.68	1.68
1.69	1.69	1.71	1.72	1.73	1.73	1.75	1.76	1.80	1.81
1.83	1.83	1.83	1.86	1.86	1.86	1.89	1.90	1.96	2.00
2.02	2.18	2.29	2.34						

Araştırmacının iddiasının geçerli olup olmadığını %95 güven düzeyinde test ediniz.
Kitle ortalaması için %95 güven düzeyinde güven aralığını oluşturunuz.

Çözüm:

$1-\alpha=95\ (güven\,düzeyi)$

$\alpha=0.05\quad(anlam\,düzeyi)$

$\sigma^2=0.05\ mg$ 'dir.

$n=84$ veriden $\bar{x}=1.603$ olarak hesaplanır.

Kitle varyansı σ^2 biliniyor.

1) Hipotez kurulur

$H_0:\mu=1.58$

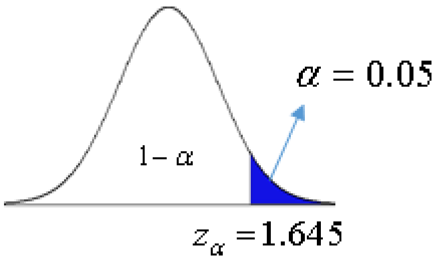
$H_1:\mu>1.58$

2) Test istatistiği:

$$z_t=\frac{\bar{x}-\mu}{\sigma/\sqrt{n}}=\frac{1.603-1.58}{0.224/\sqrt{84}}=0.94$$

olarak hesaplanır.

3) Karar aşaması:



(z_α : α 'ya bağlı olarak standart normal dağılım tablosundan bulunur.)

$z_t=0.94<z_\alpha=1.645$ olduğunda H_0 hipotezi red edilemez.

Yorum: : $100\,cm^3$ idrardaki keratin miktarının ortalamasının 1.58 mg olduğu %95 güven düzeyinde söylenebilir.

- p değeri hesaplanarak hipotez testi yapılabilir.

$$\begin{aligned}p&=P(Z\geq z_t)\\&=P(Z\geq 0.94)\\&=1-P(Z<0.94)\\&=0.1736\end{aligned}$$

$p=0.1736>\alpha=0.05$ olduğunda H_0 hipotezi red edilemez.

- Güven aralığı;

$$1-\alpha = 0.95 \Rightarrow \alpha = 0.05 \Rightarrow \frac{\alpha}{2} = 0.025$$

$z_{\alpha/2} = z_{0.025} = 1.96 \rightarrow$ standart normal dağılım tablo değeri

$$P(\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}) = 1-\alpha$$

$$\hat{\theta}_1 = \bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} = 1.603 - 1.96 \times \frac{0.224}{\sqrt{5}} = 1.555$$

$$\hat{\theta}_2 = \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}} = 1.603 + 1.96 \times \frac{0.224}{\sqrt{5}} = 1.651$$

%95 güven düzeyinde kitle ortalamasını içeren aralık (1.555; 1.651) ‘dir.

KAYNAKLAR

1. Uygulamalı İstatistik (1994)
   Ayşen APAYDIN , Alaettin KUTSAL, Cemal ATAKAN
2. Olasılık ve İstatistik Problemler ve Çözümleri ile (2008)
   Prof. Dr. Semra ERBAŞ
3. Olasılık ve İstatistik (2006)
   Prof. Dr. Fikri Akdeniz
4. Olasılık ve İstatistiğe Giriş I-II (2011)
   Prof. Dr. Fikri Öztürk
5. Fikri Öztürk web sitesi
   http://80.251.40.59/science.ankara.edu.tr/ozturk/index.html