IST108
OLASILIK VE İSTATİSTİK

HİPOTEZ TESTİ - 2

İçerik

Varyansın Bilinmediği Durum - t Testi

Önceki sunuda varyansın bilindiği, fakat beklentinin bilinmediği durumlar incelendi.

Hem beklentinin hem de varyansın bilinmediği durumlarda **t testi** uygulanır.

Hem beklentinin hem de varyansın bilinmediği durumda aşağıdaki sıfır hipotezi (H_0), alternatif hipotezine (H_1) karşı test edelim.

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

Burada σ^2 bilinmediği için sıfır hipotezi artık basit hipotez değildir.

Varyansın Bilinmediği Durum - t Testi

Önceki gibi sıfır hipotezini, örnek ortalaması $(\bar{X})$, μ_0'dan çok uzakta olduğunda reddetmek anlamlıdır.

Fakat varyansın bilinmediği durumda örnek ortalamasının μ_0'dan ne kadar uzakta olabileceği ise örnekten hesaplanacak varyansa bağlıdır.

Varyansın Bilinmediği Durum - t Testi

Artık varyans da bir bilinmeyen olduğu için, varyansı aşağıdaki örnek varyansı ile tahmin edebiliriz.

$$S^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}$$

Bu durumda H_0'ı aşağıdaki T değeri büyük olduğunda reddetmek anlamlı olur.

$$T = \frac{\sqrt{n}}{S}(\bar{X} - \mu_0)$$

Varyansın Bilinmediği Durum - t Testi

Hipotezi reddetmek için T'nin ne kadar büyük olacağını belirlemek için H_0 doğru iken T'nin dağılımına bakmalıyız.

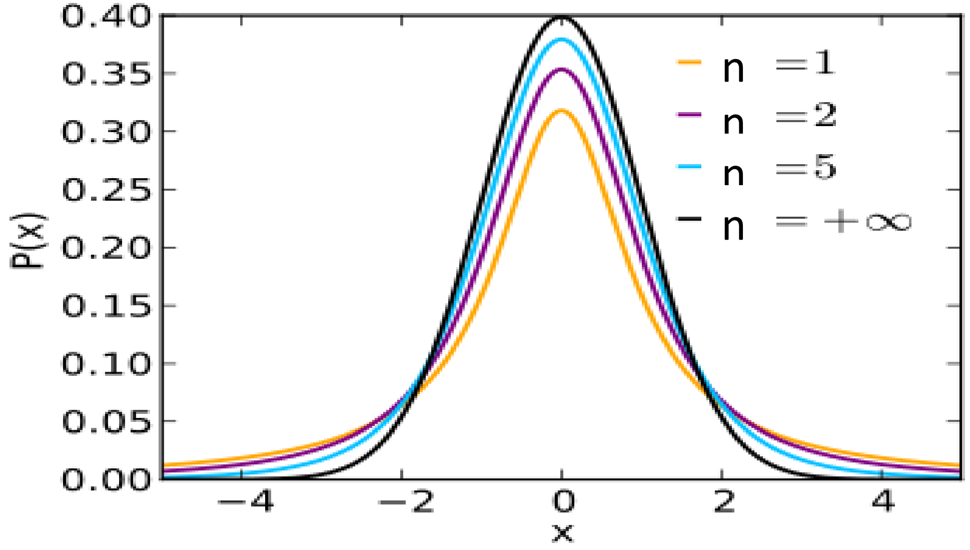
T istatistiği, $\mu = \mu_0$ olduğunda serbestlik derecesi $n - 1$ olan bir **t dağılımına** sahiptir.

t Dağılımı

Serbestlik derecesi n olan bir t dağılımı, sıfır etrafında simetrik ve standart normal dağılıma benzeyen bir eğriye sahiptir.

$$P\left(T_n \geq t_{\alpha,n}\right) = \alpha$$

Farklı α ve n değerleri için $t_{\alpha,n}$ değerlerini gösteren tablolar mevcuttur.



Bu şekil wikipedia.org'dan alınmıştır.

Varyansın Bilinmediği Durum - t Testi

Çift Yanlı Hipotez Testi

$$P\big((T < -c) \cup (T > c)\big) = 2 \times P(T > c) = \alpha \;\rightarrow P(T > c) = \alpha/2$$

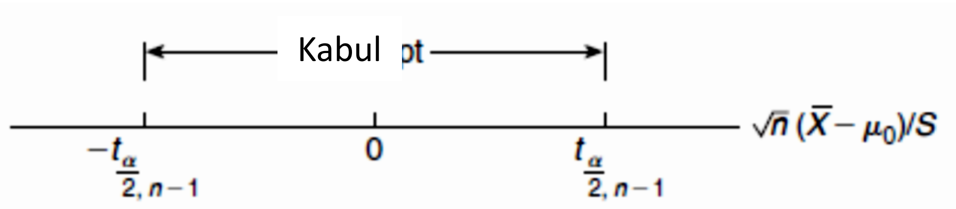
$$c = t_{\alpha/2,n-1}$$

Varyansın Bilinmediği Durum - Çift Yanlı t Testi

$H_0: \mu = \mu_0$

$H_1: \mu \neq \mu_0$

- $\frac{\alpha}{2}$ değerini kullanarak
- t tablosunda $t_{\frac{\alpha}{2},n-1}$'i bul.
- T'yi bul. $T = \frac{\sqrt{n}}{S}(\bar{X} - \mu_0)$



Daha sonra H_0'ı;

- kabul et, eğer $-t_{\alpha/2,n-1} \leq T \leq t_{\alpha/2,n-1}$ ise
- reddet, eğer $T > t_{\alpha/2,n-1}$ veya $T < -t_{\alpha/2,n-1}$ ise

t Tablosu

Örneğin, 10 elemanlı bir örnekte $\alpha = 0,05$ önem seviyesine karşılık gelen $t_{\alpha/2,n-1} = t_{0,025,9} = 2,262$ olur.

	α											
n	0.25	0.2	0.15	0.1	0.05	0.025	0.02	0.01	0.005	0.0025	0.001	0.0005
1	1.000	1.376	1.963	3.078	6.314	12.710	15.890	31.820	63.660	127.300	318.300	636.600
2	0.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.090	22.330	31.600
3	0.765	0.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841	7.453	10.210	12.920
4	0.741	0.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604	5.598	7.173	8.610
5	0.727	0.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773	5.893	6.869
6	0.718	0.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.208	5.959
7	0.711	0.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029	4.785	5.408
8	0.706	0.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501	5.041
9	0.703	0.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.297	4.781

Örnek 1

Bir kliniğin, kolesterol seviyesi orta ve yüksek seviye olan hastalar arasından kolesterolü düşüren yeni bir ilacı test etmek için gönüllüler seçiliyor. 50 gönüllüye 1 ay boyunca ilaç veriliyor ve kolesterol değişimi gözleniyor. Eğer kolesterol düşümü ortalama 14,8 ve örnek standart sapması 6,4 ise bundan nasıl bir sonuç çıkartırız? 0,05 önem seviyesi için hesaplayarak inceleyiniz.

Örnek 1

Kolesterol düşümü olup olmadığını test edelim.

- $H_0: \mu = 0$
- $H_1: \mu \neq 0$

$t_{\alpha/2,n-1} = t_{0,025,49} \cong 2,009$

$$T = \frac{\sqrt{n}}{s}(\bar{X} - \mu_0) = \frac{\sqrt{50}}{6,4}(14,8 - 0) = 16,352$$

$16,352 > 2,009$ olduğu için sıfır hipotezi reddedilir. Karşı hipotez kabul edilir yani kolesterol düşümü olduğu söylenir.

Fakat hastaların kolesterol seviyelerinde meydana gelen düşmenin ilaç kaynaklı olup olmadığı kanıtlanmış değildir.

Örnek 2

Halk sağlığı müdürlüğü ortalama musluk suyu kullanımının ev başına günlük 350 litre olduğunu iddia etmektedir. Bu iddiayı doğrulamak için 20 ev rastgele seçilmiş ve günlük su kullanımları aşağıdaki gibi kaydedilmiştir.

340	344	362	375	318
356	386	354	364	360
332	402	340	355	338
362	322	372	324	370

Bu veri iddia ile tutarlı mıdır?

Örnek 2

340	344	362	375	318
356	386	354	364	360
332	402	340	355	338
362	322	372	324	370

Verilerin ortalaması $353,8$ ve standart sapması $21,8478$

Örnek 2

Önem seviyesi $\alpha = 0{,}1$ olsun (Hata Tipi I olasılığı $\%10$ olsun).

- $H_0\colon \mu = 350$
- $H_1\colon \mu \neq 350$

$t_{\alpha/2,n-1} = t_{0{,}05,\,19} = 1{,}73$

$T = \frac{\sqrt{n}}{s}(\bar{X} - \mu_0) = \frac{\sqrt{20}}{21{,}8478}(353{,}8 - 350) = 0{,}7778$

$-1{,}730 < 0{,}7778 < 1{,}730$ olduğu için sıfır hipotezi kabul edilir.

Veriler yetkilinin iddiası ile tutarlıdır.

Varyansın Bilinmediği Durum - Tek Yanlı t Testi

Hipotez test problemi:

- $H_0: \mu = \mu_0$ (veya $\mu \leq \mu_0$)
- $H_1: \mu > \mu_0$

Bu hipotez testi, varyans bilindiğindeki tek yanlı teste benzer.

- α değerini kullanarak
- t tablosunda $t_{\alpha,n-1}$'i bul.
- T'yi bul. $T = \frac{\sqrt{n}}{s}(\bar{X} - \mu_0)$

Daha sonra H_0'ı;

- kabul et, eğer $T \leq t_{\alpha,n-1}$ ise
- reddet, eğer $T > t_{\alpha,n-1}$ ise

Varyansın Bilinmediği Durum - Tek Yanlı t Testi

Hipotez test problemi:

- $H_0: \mu = \mu_0$ (veya $\mu \geq \mu_0$)
- $H_1: \mu < \mu_0$

- α değerini kullanarak
- t tablosunda $t_{\alpha,n-1}$ ‘ i bul.
- T’yi bul. $T = \frac{\sqrt{n}}{s}(\bar{X} - \mu_0)$

Daha sonra H_0’ı;

- kabul et, eğer $T \geq -t_{\alpha,n-1}$ ise
- reddet, eğer $T < -t_{\alpha,n-1}$ ise

Örnek 3

Yeni bir fiberglass lastik üreticisi lastiklerin ortalama ömrünün en az 40.000km'den fazla olduğunu iddia etmektedir. Bu iddiayı doğrulamak için 12 adet lastik test edilmiş ve yaşam süreleri aşağıdaki gibi kaydedilmiştir. (Birim x1000km).

Lastik	1	2	3	4	5	6	7	8	9	10	11	12
Yaşam Süresi	36,1	40,2	33,8	38,5	42	35,8	37	41	36,8	37,2	33	36

Üreticinin iddiasını %5 önem seviyesine göre test edin.

Örnek 3

Lastik	1	2	3	4	5	6	7	8	9	10	11	12
Yaşam Süresi	36,1	40,2	33,8	38,5	42	35,8	37	41	36,8	37,2	33	36

Bu verilerin ortalaması 37,2833 ve standart sapması 2,7319.

Örnek 3

Önem seviyesi : %5

$H_0: \mu \leq 40$

$H_1: \mu > 40$

$t_{\alpha,n-1} = t_{0{,}05,\,11} = 1{,}796$

$$T = \frac{\sqrt{n}}{s}(\bar{X} - \mu_0) = \frac{\sqrt{12}}{2{,}7319}(37{,}2833 - 40) = -3{,}4448$$

$-3{,}4448 < 1{,}796$ olduğundan sıfır hipotezi kabul edilir.

Veriler üreticinin iddiası ile tutarsızdır.

Özet

Özet

Varyansın Bilinmediği Durum - t Testi		
$S^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}$		$T = \frac{\sqrt{n}}{S}(\bar{X} - \mu_0)$
Çift Yanlı t Testi H_0: $\mu = \mu_0$ H_1: $\mu \neq \mu_0$	Tek Yanlı t Testi H_0: $\mu = \mu_0$ veya $(\mu \leq \mu_0)$ H_1: $\mu > \mu_0$	Tek Yanlı t Testi H_0: $\mu = \mu_0$ veya $(\mu \geq \mu_0)$ H_1: $\mu < \mu_0$
H_0 kabul: $-t_{\alpha/2,n-1} \leq T \leq t_{\alpha/2,n-1}$	H_0 kabul: $T \leq t_{\alpha,n-1}$	H_0 kabul: $T \geq -t_{\alpha,n-1}$
H_0 red: $T > t_{\alpha/2,n-1}$ veya $Z < -t_{\alpha/2,n-1}$	H_0 red: $T > t_{\alpha,n-1}$	H_0 red: $T < -t_{\alpha,n-1}$

İki Normal Yığının Ortalamalarının Eşitlik Testi

Bazen iki farklı yaklaşımın aynı sonucu verip vermediğine karar vermek isteriz. Bu durum iki normal yığının aynı ortalama değere sahip olması hipotezini test ederiz.

$X_1, X_2, \ldots, X_n$ ortalaması (μ_x) bilinmeyen ama varyansı (σ_x^2) bilinen bir yığından n elemanlı ve $Y_1, Y_2, \ldots, Y_m$ ortalaması (μ_y) bilinmeyen ama varyansı (σ_y^2) bilinen başka bir yığından m elemanlı rastgele seçilen örnekler olsun.

İki Normal Yığının Ortalamalarının Eşitlik Testi

X değerlerinin örnekleme ortalaması, μ_x' i ve Y değerlerinin örnekleme ortalaması, μ_y'yi tahmin etmek için kullanılabileceğinden bu ortalamaların farkı, $\mu_x - \mu_y$ *' yi* tahmin etmek için kullanılabilir. Bu durumda $H_0 : \mu_x - \mu_y = 0$ yazılabilir.

Hipotez test problemi:

- $H_0 : \mu_x = \mu_y$ *veya* $\mu_x - \mu_y = 0$
- $H_1 : \mu_x \neq \mu_y$ *veya* $\mu_x - \mu_y \neq 0$

Bu durumda H_0'ı;

- kabul et, eğer $-c \leq (\bar{X} - \bar{Y}) \leq c$ ise
- reddet, eğer $(\bar{X} - \bar{Y}) < -c$ veya $(\bar{X} - \bar{Y}) > c$ ise

İki Normal Yığının Ortalamalarının Eşitlik Testi

α anlam düzeyinde c değerini belirlemek için örnek ortalamalarının birbirlerinden uzaklığının hipotez doğru iken dağılımına bakmalıyız.

$$\bar{X} - \bar{Y} \sim \mathcal{N}\left(\mu_x - \mu_y, \frac{\sigma_x^2}{n} + \frac{\sigma_y^2}{m}\right)$$

$$P\big((\bar{X} - \bar{Y} < -c) \cup (\bar{X} - \bar{Y} > c)\big) = P\left(\left(Z < \frac{-c-(\mu_x-\mu_y)}{\sqrt{\frac{\sigma_x^2}{n}+\frac{\sigma_y^2}{m}}}\right) \cup \left(Z > \frac{c-(\mu_x-\mu_y)}{\sqrt{\frac{\sigma_x^2}{n}+\frac{\sigma_y^2}{m}}}\right)\right) = \alpha$$

Eğer H_0 doğru ise ($\mu_x - \mu_y = 0$):

$$P\left(\left(Z < \frac{-c}{\sqrt{\frac{\sigma_x^2}{n}+\frac{\sigma_y^2}{m}}}\right) \cup \left(Z > \frac{c}{\sqrt{\frac{\sigma_x^2}{n}+\frac{\sigma_y^2}{m}}}\right)\right) = 2 \times P\left(Z > \frac{c}{\sqrt{\frac{\sigma_x^2}{n}+\frac{\sigma_y^2}{m}}}\right) = \alpha \to P\left(Z > \frac{c}{\sqrt{\frac{\sigma_x^2}{n}+\frac{\sigma_y^2}{m}}}\right) = \alpha/2$$

$$c = z_{\alpha/2}\sqrt{\frac{\sigma_x^2}{n} + \frac{\sigma_y^2}{m}}$$

İki Normal Yığının Ortalamalarının Eşitlik Testi

Hipotez test problemi:

- $H_0 : \mu_x = \mu_y$
- $H_1 : \mu_x \neq \mu_y$

1. $1 - \frac{\alpha}{2}$ değerini kullanarak
2. Standart Normal Dağılım tablosunda $z_{\alpha/2}$ ‘ yi bul.
3. Z’yi bul. $Z = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_x^2}{n} + \frac{\sigma_y^2}{m}}}$

Yani H_0’ı;

- kabul et, eğer $-z_{\alpha/2} \leq Z \leq z_{\alpha/2}$ ise
- reddet, eğer $Z > z_{\alpha/2}$ veya $Z < -z_{\alpha/2}$ ise

İki Normal Yığının Ortalamalarının Eşitlik Testi Tek-yanlı

Hipotez test problemi:

- $H_0 : \mu_x = \mu_y$ $(\mu_x \leq \mu_y)$
- $H_1 : \mu_x > \mu_y$

1. $1 - \alpha$ değerini kullanarak
2. Standart Normal Dağılım tablosunda z_α ‘ yi bul.
3. Z’yi bul. $Z = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_x^2}{n} + \frac{\sigma_y^2}{m}}}$

Yani H_0’ı;

- kabul et, eğer $Z \leq z_\alpha$ ise
- reddet, eğer $Z > z_\alpha$ ise

Örnek 4

Araç lastiği üretmek için iki yeni metot geliştirilmiştir. Hangisini daha iyi olduğuna karar vermek için bir lastik üreticisi birinci metodu kullanarak 10 lastik ve ikinci metodu kullanarak 8 lastik üretmiştir.

Birinci metotla üretilen lastikler test edilmiş ve ömür ortalamaları 61550 km olarak hesaplanmıştır. Yaşam sürelerinin standart sapması 4000 km olan bir Normal dağılımla ifade edildiği biliniyor.

İkinci metotla üretilen lastikler ise test edildiğinde ömür ortalamaları 60025 km olarak hesaplanmıştır. Yaşam sürelerinin standart sapması 6000 km olan bir Normal dağılımla ifade edildiği biliniyor.

Üretici bu testler sonucunda iki metodun eşdeğer olduğunu düşünüyorsa, üreticinin bu iddiasını %5 önem seviyesi için test edin?

Örnek 4

Araç lastiği üretmek için iki yeni metot geliştirilmiştir. Hangisini daha iyi olduğuna karar vermek için bir lastik üreticisi birinci metodu kullanarak 10 lastik ve ikinci metodu kullanarak 8 lastik üretmiştir. (1000km birim ile)

Birinci metot ortalama yaşam süresi 61,55, standart sapma 4.

İkinci metotla ortalama yaşam süresi 60,025 standart sapması 6.

Önem seviyesi %5.

- $H_0 : \mu_x = \mu_y$
- $H_1 : \mu_x \neq \mu_y$

$$z_{\alpha/2} = z_{0,025} = 1,96$$

$$Z = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_x^2}{n} + \frac{\sigma_y^2}{m}}} = \frac{61,55 - 60,025}{\sqrt{\frac{16}{10} + \frac{36}{8}}} = 0,6175$$

- -1,96 < 0,6175 < 1,96 için sıfır hipotezi kabul edilir.