

Daha ileriye... En iyiye...

İki örneklemlili Hipotez Testleri

Bağımlı Gruplarda t testi



www.hacettepe.edu.tr

Dr. Kübra Atalay
Kabasakal

TEK ÖRNEKLEM HİPOTEZ TESTLERİ

Tek Örneklem z testi	Tek Örneklem t testi
Evren ortalaması(μ) ve evren standard sapması (σ) bilinir.	Evren ortalaması (μ) bilinir. Evren standart sapması (σ) bilinmez
Standart hata (örneklem ortalamaları dağılımının standart sapması) aşağıdaki gibi hesaplanır: $\sigma_{\bar{X}} = \sigma / \sqrt{N}$	Standart hata örneklemin standard sapması kullanılarak aşağıdaki gibi kestirilir: $S_{\bar{X}} = s / \sqrt{N}$
Örneklem ortalaması için z istatistiği: $z = \frac{\bar{X} - \mu}{\sigma_{\bar{X}}}$	Örneklem ortalaması için t istatistiği: $t = \frac{\bar{X} - \mu}{S_{\bar{X}}}$
Kritik değerler birim normal dağılım tablosun elde edilir.	Kritik değerler belirli bir serbestlik derecesindeki (N-1) t dağılımından elde edilir. Serbestlik derecesi arttıkça t değeri z değerine yaklaşır

Not: Bazı kaynaklarda örneklem ortalaması sembolü \bar{X} bazılarında da μ olarak kullanılır.

TEKRARLI ÖLÇÜMLER DESENİ

- Aynı bireylerin birden fazla işlem durumunda gözlenerek, yani bir bağımsız değişkenin tüm durumlarında ölçüm yapılarak gerçekleştirilen çalışmalar tekrarlı ölçümler deseni kullanılarak yapılan çalışmalardır.
- Bu desende aynı katılımcıların bir program öncesinde ve sonrasında ölçümleri alınarak tekrarlı ölçümler elde edilir.
- Test tekrar test deseni de denir.
- Uygulama durumlarında aynı bireyler kullanıldığı için, bir gruptaki katılımcıların diğer gruptaki katılımcılardan farklı olma riski yoktur.
- Grup içi desenler de denir.

EŞLEŞTİRİLMİŞ GRUPLAR DESENİ

- Araştırmacılar bazen iki farklı örneklem kullanırlar; ancak örneklemi oluşturulan bireyler, çalışma için önemli olduğu düşünülen ve araştırmacının kontrol etmek istediği bir veya birkaç değişkene göre benzer özelliklere sahip olacak şekilde eşleştirilirler. Böyle çalışmalar eşleştirilmiş gruplar deseni kullanılarak yapılan çalışmalardır.
- Örneğin, bireylerin sözel yeteneklerinin incelendiği bir çalışmada araştırmacı örneklemi IQ puanı ve cinsiyete göre eşleştirilebilecek bireylerle oluşturmak isteyebilir. Yani bir örnekleme IQ puanı 120 olan bir erkek varsa, diğer örnekleme de IQ puanı 120 olan bir erkek bulunacaktır.
- Tekrarlı ölçümler ve eşleştirme yöntemlerinde veri setleri birbiri ile ilişkili örneklemelerden oluştuğu için, bağımlı (ilişkili) örneklemelerle yapılan araştırmalar şeklinde tanımlanır.

BAĞIMLI ÖRNEKLEMLER/TEKARLI ÖLÇÜMLER İÇİN T İSTATİSTİĞİ

- Grup içi desenlerde iki ortalamamanın karşılaştırılmasında kullanılır.
- Her bir katılımcı için fark puanı hesaplanır.
- Hesaplanan fark puanlarının ortalamalarının 0'dan anlamlı düzeyde farklı olup olmadığı incelenir.
- Dolayısıyla burada yeni bir dağılım olan fark puanlarının dağılımı ile çalışırız.

Örnek: Araştırmacılar A ilacının bireylerin reaksiyonda bulunma (tepki verme) süresine etkisini incelemektedirler. Bu kapsamda beş katılımcının tepki verme süreleri ilacı almadan önce ve ilacı aldıktan bir saat sonra ölçülmüş ve bu ölçümler karşılaştırılmıştır.

Araştırma sorusu: A ilacı bireylerin tepki verme sürelerini etkilemekte midir?

Aşağıdaki tabloda bireylerin ilacı almadan önceki ve aldıktan yarım saat sonraki reaksiyon süreleri görülmektedir. (Fark puanları: $D=X_2-X_1$)

Katılımcı	İlacı almadan önce (X_1)	İlacı aldıktan sonra (X_2)	Fark D
A	215	210	-5
B	221	242	21
C	196	219	23
D	203	228	25
$\bar{X}_D = (\sum D)/n = 64/4 = 16$			$\sum D = 64$

Hipotez testinin aşamaları:

1. Evren dağılım özelliklerinin belirlenmesi
2. Hipotezlerin kurulması
3. Karar kuralının belirlenmesi
4. Örneklem istatistiğinin hesaplanması
5. Karar verme
6. Sonuçların yorumlanması

1. Evren dağılım özelliklerinin belirlenmesi

- Test tekrar test deseni dolayısıyla bağımlı örneklemeler
- İki uygulama : ilaç öncesi ve ilaç sonrası
- Bağımlı değişken: Tepki süresi
- Bağımsız değişken: İlaç
- Bağımlı değişken sürekli ve en az eşit aralık ölçeğinde.
- Normal dağılım

2. Hipotezlerin Kurulması

Araştırmacılar, iki uygulama arasında evrende bir fark olup olmadığı ile, yani fark puanları ile ilgilenirler.

Evrendeki fark puanlarının ortalaması: μ_D

$$H_0: \mu_D=0$$

Sıfır hipotezine göre, tüm evrendeki fark puanlarının ortalaması sıfırdır. Bazı bireyler pozitif fark puanlarına bazıları ise negatif fark puanlarına sahip olabilir; ancak bu farklar şans eseri oluşmuş, sistematik olmayan farklardır ve ortalamaları sıfırdır.

$$H_1: \mu_D \neq 0$$

Uygulamalardan birine ait puanlar sistematik bir şekilde diğer uygulamaya ait puanlardan yüksektir (düşüktür).

3. Karar kuralının belirlenmesi

- *Alfa değerinin seçilmesi:* Alfa değeri 0.05 olsun
- *Kritik t değerlerinin belirlenmesi:* Testimiz yönsüz olduğundan t tablosunda her iki uçtaki toplam alanın 0.05 (veya her bir uçtaki alanın 0.025 olduğu sütun) olduğu sütun ile Serbestlik derecesi = $4-1=3$ (Toplamda 4 fark puanımız var bu nedenle $N=4$) olan satırda yer alan t değeri okunur. Bu değer 3.183'tür.
- Kritik t değerleri -3.183 ve 3.183 dir.
- Karar kuralı:
 - Örneklemden elde edilecek olan t istatistiği kritik değer olan 3.183'den büyük veya - 3.183'den küçük olursa H_0 reddedilir.
 - Örneklemden elde edilecek olan t istatistiği - 3.183 ile 3.183 aralığında olursa H_0 kabul edilir.

4. Örneklem istatistiğinin hesaplanması

- Bağımlı örneklemelerde veri setini, X değerleri yerine fark puanları, yani D değerleri oluşturur. Bu sebeple ilgilenilen evren ortalaması fark puanlarının ortalamasıdır ($\mu_D=?$).

$$t = \frac{\text{Örneklem İstatistiği} - \text{hipotez edilen evren parametresi}}{\text{Kestirilen standart hata}}$$

$$t = \frac{\bar{X}_D - \mu_D}{S_{\bar{X}_D}}$$

Standart hatanın hesaplanması:

- $SH = s_{M_D} = \frac{s}{\sqrt{N}}$

- s: fark puanlarının standart sapmasıdır, $s = \sqrt{\frac{KT(Kareler\ toplamı)}{sd(Serbestlik\ derecesi)}}$

$$KT = \sum (D - M_D)^2 = (-5 - 16)^2 + (21 - 16)^2 + (23 - 16)^2 + (25 - 16)^2 = 441 + 25 + 49 + 81 = 596$$

$$s = \sqrt{\frac{KT}{sd}} = \sqrt{\frac{596}{3}} = 14.09$$

$$SH = s_{M_D} = \frac{s}{\sqrt{N}} = \frac{14.09}{\sqrt{4}} = 7.005$$

t istatistiği

$$t = \frac{\bar{X} - \mu_D}{S_{M_D}}$$

$$t = \frac{16 - 0}{7.005}$$

$$t = 2.28$$

5. Karar verme

Elde ettiğimiz test istatistiğimiz 2.28 kritik değer olan 3.183'den küçük olduğu için H_0 red edilmez.

6. Yorumlama

A ilacının bireylerin tepki verme süreleri üzerinde anlamlı bir etkisi yoktur ($t = 2.28, p > 0.05$)

Etki Büyüklüğü

- Not: Anlamlı fark bulamadığımız durumlarda etki büyüklüğü raporlamayız.
- Etki büyüklüğünü fark puanları için hesaplarız.

$$d = \frac{\text{Örneklem farklarının ortalaması}}{\text{farkların standart sapması}} = \frac{\bar{X}_D}{s}$$

$$d = \frac{16}{14.09} = 1.12$$

Güven Aralığı: Ortalama fark değeri için

- $\mu_D = \bar{X}_D \pm (SH * t)$
- μ_D için % 95 güven aralığı $16 \pm (7 * 3.183) = 16 \pm (22)$
- %95 olasılıkla evrende farkların ortalaması -6 ile 38 aralığındadır.

Not: İstatistiksel olarak anlamlı bir fark bulamadığımızda güven aralığı 0'ı içerir.

Varsayımlar

Bağımsız gözlem: Her bir uygulama durumu içerisindeki gözlemlerin bağımsız olması gerekmektedir.

Normal dağılım: Evrendeki fark puanlarının (D değerleri) dağılımı normal olmalıdır. Bu varsayım özellikle küçük örneklemeler için önemlidir. $n > 30$ ise bu varsayım sağlanmış olur.