

İki Bağımlı Grup Ortalaması için Hipotez Testi (Paired t Test)

İki grubun ortalamalarının karşılaştırılmasında, grupların bağımlı olması durumunda kovaryans bilgisi artık sıfır olmayacaktır. Grupların bağımlı olması için genel olarak iki yaklaşımdan söz edilebilir: Eşleştirme (Paired) ve İşlem Öncesi-İşlem Sonrası.

Eşleştirme tasarımı bir örnek üzerinden anlatılacaktır. Şimdi korona virüsüne yakalanan insanlara verilen A ve B ilaçlarının etkisini incelemek istediğimizi, ancak ilaçların başarısının hastanın bazı klinik karakteristiklerine (sigara içme, yaş, cinsiyet vb.) göre farklılaştığını düşünelim. Bu durumda klinik karakteristiklerin etkisini ortadan kaldırarak sadece ilaçların performansına odaklanmak için aynı ya da benzer klinik karakteristiklere sahip ikişerli kişilerden oluşan n çift oluşturmalı ve daha sonra her bir çiftte bulunan kişilerden birisini 1. gruba ve diğerini de 2. gruba atamalıyız. Böylece klinik göstergelerin etkisi arındırılmış olacak ve elimizdeki $2n$ sayıda bireyin n tanesi 1. grupta, diğer n tanesi de 2. grupta yer alacaktır. İşte bu şekilde yapılan bir tasarım sonucunda elde edilen gruplar bağımlı olacaktır.

İşlem öncesi-işlem sonrası tasarımda ise aynı bireyler üzerinde bir sürecin başında ve sürecin sonunda yapılan ölçümler neticesinde elde edilen ortalamalar arası farklılık ile ilgilenilmektedir. Örneğin “Diyet programından önceki ve sonraki ortalama ağırlıklar arasında fark var mıdır?”, “Yabancı dil kursundan önce ve sonra ortalama YDS puanları arasında fark var mıdır?” vb. sorular ile karşılaşıldığında, aynı kişiler üzerinde ölçüm yapıldığı için veri setinin bağımlı olduğuna karar verilecektir. Genellikle bu ders boyunca işlem öncesi-işlem sonrası tasarımı üzerinde durulacaktır.

Bağımlı iki grubun ortalamaları için ileri sürülebilecek hipotezler aşağıdaki gibidir.

$$\begin{array}{lll} H_0: \mu_{\text{önce}} = \mu_{\text{sonra}} & H_0: \mu_{\text{önce}} = \mu_{\text{sonra}} & H_0: \mu_{\text{önce}} = \mu_{\text{sonra}} \\ H_1: \mu_{\text{önce}} > \mu_{\text{sonra}} & H_1: \mu_{\text{önce}} < \mu_{\text{sonra}} & H_1: \mu_{\text{önce}} \neq \mu_{\text{sonra}} \end{array} \quad (9.6)$$

Burada $\mu_{\text{önce}}$ işlemden önceki bilinmeyen ortalama parametresini, μ_{sonra} ise işlemden sonraki bilinmeyen ortalama parametresini ifade etmektedir. (9.6) ile verilen hipotezlerin (9.7)’de verilen hipotezlerle eşdeğer olduğu aşıkardır.

$$\begin{array}{lll} H_0: \mu_d = 0 & H_0: \mu_d = 0 & H_0: \mu_d = 0 \\ H_1: \mu_d > 0 & H_1: \mu_d < 0 & H_1: \mu_d \neq 0 \end{array} \quad (9.7)$$

Burada $\mu_d = \mu_{\text{önce}} - \mu_{\text{sonra}}$. (9.7) ile verilen hipotezlerden de anlaşılacağı gibi, söz konusu fark işleminden sonra problem tek kitle ortalama parametresi için hipotez testine dönmektedir. O halde tek kitle durumunda kullanılan t testinde X_i yerine $d_i = X_i(\text{önce}) - X_i(\text{sonra})$, $\bar{X}$ yerine $\bar{d}$ ve $s_{\bar{X}}$ yerine de $s_{\bar{d}}$ kullanılacaktır. Bu değerler aşağıdaki gibi hesaplanır.

$$\bar{d} = \frac{1}{n} \sum_{i=1}^n d_i, s^2 = \frac{1}{n-1} \sum_{i=1}^n (d_i - \bar{d})^2, s_{\bar{d}}^2 = \frac{s^2}{n}$$

O halde verinin normal olduđu varsayımı altında (9.7) ile verilen hipotezlerin testinde kullanılacak test istatistiđi,

$$t = \frac{\bar{d} - \mu_d}{s_{\bar{d}}} \sim t_{n-1} \quad (9.8)$$

olarak ifade edilir. Burada n değeri eşleştirme tasarımında çiftlerin sayısıyken, işlem öncesi- işlem sonrası tasarımında gözlem sayısıdır. Test işlemi n-1 serbestlik dereceli t dağılımı üzerinde ret ve kabul bölgeleri oluşturularak ya da p değeri kullanılarak gerçekleştirilir.