

## TEKNOLOJİ VE ENDÜSTRİ ALANLARINDA VERİ ANALİZİ ÜZERİNE BİR UYGULAMA

Dr. Fazıl DEMİRCİ\*

### GİRİŞ

İstatistik ilmi veya istatistik metodları, hangi alanda olursa olsun sayısal değerlerin sistematik olarak toplanması ve prezente edilmesini gerçekleştiren ve bu değerleri çeşitli yollarla karşılaştıran, özetleyen ve analiz ederek popülasyona genelleştiren bir bilim dalıdır. Popülasyon'lar hakkında geniş bilgileri elde edebilmek için örnekleme yolları kullanılarak istenen bir kütle hakkında gerekli bilgiler toplanır ve bu bilgilerin anılan popülasyonları temsil edip etmeyecekleri de kurulacak olan hipotezlerle kontrol edilir.

### İstatistik'in Ana Bölümleri

İstatistik ilmi iki ana bölümde açıklanır :

1 — Teorik İstatistik : Doğal olarak formüllerle açıklanan matematik kökenlidir. Örnek olarak, toplama metodları, varyasyon çalışmalarları ve ilişkileri hesaplama metodları, standart sapma, korelasyon katsayısı vs.

2 — Uygulamalı İstatistik : İstatistik metodlarının uygulanmalarıyla ortaya çıkan bir bölüm olup aşağıdaki alanlarda uygulanabilmektedir.

- a — Biyometri
- b — Psikometri
- c — Ekonomi İstatistiği
- d — Hayat İstatistiği

(\*) A.Ü. Ziraat Fakültesi Biyometri Anabilim Dalı.

Bunların dışında, istatistik metodları hemen hemen her konuda yani her bilim dalında kullanma olanağı bulabilmiştir.

Endüstri alanında, yeni teknolojik gelişmelere paralel olarak en önemli hususun rekabeti teşvik eden kalite dereceleri ve aynı malı üreten fabrikalardaki kalite durumlarıdır.

Tüketici açısından konuya bakılacak olursa, fiyatın ucuzluğu ve malın dayanıklılığı üzerinde durulduğu görülür. Fiyatların aynı olduğu ürünlerde ise dayanıklılık bir tercih sebebi olarak görülür. Burada önemli bir husus ortaya çıkmış olur ki bu da aynı malı üreten fabrikaların ürünlerinin tercih sebepleri istatistiki olarak kontrol edilmesi durumlarıdır. Yeni teknolojilerin gelişmesiyle tüketicilerin korunmaları için tedbirlerin devlet tarafından alınması ve bu teknolojinin bazı üretim, tüzük ve standartlarının uygulanıp uygulanmadığı kontrolü önemlilik kazanır. Bu durumlarla birlikte istatistik metodlarının bu alanlardaki yerlerinin ne kadar önemli olduğu ortaya çıkmış olur. Yukarıda belirtilen hususlara aşağıdaki örnekleri vermek mümkündür :

1 — Bir litre benzinle alınacak yolu artırabilmek için, benzene eklenen bir katkı maddesinin gerçekten yolu artırıp artırmadığının kontrol edilmesi. Yani katkı maddesi kullanan ve kullanmayan iki benzin istasyonunun rekabetinin gerçek olup olmadığının kontrol edilmesi.

2 — Yeni teknolojiyle gelmiş olan ekmeğin fabrikasında üretilmekte olan ekmeğin gramajlarının gıda maddeleri tüzüğünde veya belediyelerin öngördüğü gramajda üretilip üretilmediğinin kontrol edilmesi veya iki ayrı ekmeğin fabrikasında üretilen ekmeğin ağırlık bakımından aynı olup olmadıklarının kontrol edilmesi.

3 — Endüstride üretilen yağların tip ve çeşitlerine göre gerilme durumlarının gerçekten imaline izin verildiği standartlarda üretilip üretilmediğinin kontrol edilmesi veya yağların kanunlarda belirtilen ağırlıkları tahammül edip etmeyeceklerinin kontrol edilmeleri.

4 — Her hangi bir madeni malda kullanılan kaplama maddelerinin dayanıklılıklarının ölçülmesi ve hangi kaplama maddesinin daha dayanıklı olduğunun tesbit edilmesi.

5 — Diş hekimliğinde kullanılan dolgu maddelerinin dayanıklılık ve ete olan uyumlarının kontrol edilmesi gibi daha nice örnek verilebilir.

#### Formülasyon :

Çeşitli endüstri konularında ciddi istatistiki kontrollerin yapılabilmesi için üzerinde çalışılacak materyal tesbit edilir ve daha sonra populasyon değerinin olup olmadığına bakılır eğer populasyon varyansı biliniyorsa aşağıdaki formüller kullanılır :

$$Z = (\bar{X} - \mu) / \sigma\bar{x}$$

Populasyon varyansı bilinmiyorsa :

$$t = (\bar{X} - \mu) / S\bar{x}$$

$\bar{X}$  = Örnek Ortalaması

$\mu$  = Populasyon Ortalaması

$\sigma$  = Populasyona ait standart sapma

$S\bar{x}$  = Örnekten hesaplanan standart sapma

Varyansı belli olan iki normal populasyon ortalamaları arasındaki farkı kontrolü için :

Ortalamaları  $\mu_1$  ve  $\mu_2$  varyansları da  $\sigma_1^2$ ,  $\sigma_2^2$  olan iki populasyon varsayılarak bunlardan birincisinden  $n_1$  büyüklüğünde örnek çekilirse

$$X_{11}, X_{12}, X_{13}, \dots, X_{1n_1}$$

İkinci populasyondan da  $n_2$  büyüklüğünde örnek çekilirse :

$$X_{21}, X_{22}, X_{23}, \dots, X_{2n_2}$$

Bunların ortalamaları alınarak birinci örnek ortalaması :

$$\bar{X}_1, N (\mu_1, \sigma_1^2 / n_1)$$

ve ikinci örnek ortalaması :

$$\bar{X}_2, N (\mu_2, \sigma_2^2 / n_2)$$

şeklinde bir dağılım gösterir.

Burada görüldüğü gibi iki ayrı populasyonun ortalamaları arasındaki farkın önemli olup olmadığını kontrol edilmesi isteniyorsa öncelikle buna uygun bir hipotez kurulur ve daha sonra kurulan bu hipotez test edilir.

Hipotez :

$$H_0: \mu_1 = \mu_2$$

şeklinde kurulur ki bunun anlamı : her iki örnek, ortalaması aynı olan bir popülasyondan çekilmiştir. Veya örnekler ortalamaları aynı olan iki popülasyondan çekilmiştir. Bunlar :

$$N (\mu_1 - \mu_2, \sigma_1^2 / n_1 + \sigma_2^2 / n_2)$$

şeklinde bir dağılım gösterirler. Bunun sebebi ise bu iki örneğin birbirlerinden bağımsız olmalarıdır.

Her hangi bir hipotezin kontrol edilebilmesi için aşağıdaki işlemler yapılmalıdır :

- Yapılacak işleme göre bir hipotez ve buna karşı bir de bu hipotez red edildiğinde kabul edilmesi gereken karşıt hipotezin kurulması.
- Test için önemlilik seviyesi önceden tesbit edilir ( $\alpha$ ).
- Popülasyon varyansı biliniyorsa Z ve bilinmiyorsa t belirlenir.
- Eğer hesaplanan Z ve t değerleri tablo değerlerinden küçük h hipotezi kabul edilir büyük ise h hipotezi red edilir ve karşıt hipotez kabul edilir.

#### Uygulama :

Elektrik lambası üreten iki firmadan alınan 50 adet lamba test edilmiş ve X firmasının ürettiği lambaların ortalama 1282 saat, standart sapması 80 saat, Y firmasının lambaları ise ortalama 1208 saat ve standart sapması da 94 saat olarak ortaya çıkmıştır. Bu iki firmasının ürettiği lambaların arasındaki farkın gerçekten önemli olup olmadığını kontrol ediniz.

Hipotez :

İki firma ortalaması arasındaki fark tesadüften ileri gelmiştir. Yani

$$h_0: \mu_1 = \mu_2$$

$$h_1: \mu_1 \neq \mu_2$$

Popülasyon varyansı bilindiği için : Z testi kullanılır :

$$Z = (\bar{x}_1 - \bar{x}_2) / \sqrt{(\sigma_1^2 / n_1 + \sigma_2^2 / n_2)}$$

$$\bar{x}_1 - \bar{x}_2 = 1282 - 1208 = 74$$

$$(80)^2 / 50 = 128$$

$$(94)^2 / 50 = 176.72$$

$$\sqrt{128 + 176.72} = 17.46$$

$$Z = 74 / 17.46 = 4.24$$

Hesaplanan Z değeri  $\alpha = \% 5$  göre tablo değeri olan 1.96'dan daha büyük olduğu için h hipotezi red edilir ve bu 74 saatlık farkın tesadüfen ileri gelmediği anlaşılır. X firması Y firmasından lambaların yanma süresi bakımından daha iyi olduğu yani, X firmasının ürettiği lambalar Y firmasının ürettiği lambalardan daha fazla uzun ömürlü olduğu ortaya çıkar ve X firmasının lambaları tüketici tarafından bir tercih sebebi olur.

#### KAYNAKLAR

- Chaudhry, M., Introduction to Statistical Theory I, İlmî Kitabkhana Lahore 1973.  
Choudhry, M., Introduction to Statistical Theory II, İlmî Kitabkhana Lahore 1973.  
Demirci, F., A.Ü. Ziraat Fakültesi İstatistik Uygulamaları ders notları Ankara 1991.  
Orhan, D. ve ark., İstatistik Metodları, Ziraat Fakültesi yayınları No : 861, Ankara, 1983.  
Vedi, N., ve ark., Mebadi Evveliyeh Fil İhsa, Awiley Arabook, 1979 Toronto.