4. TAHMİN SONUÇLARININ DEĞERLENDİRİLMESİ

4.1. Katsayıların Yorumu

$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \ldots + \beta_k X_{ki} + u_i$ gibi çok açıklayıcı değişkene sahip bir modelde,

anakütle regresyon fonksiyonu, $E(Y_i|X_{1i}, X_{2i}, \ldots, X_{ki}) = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \ldots + \beta_k X_{ki}$ dir. Böyle bir modelde katsayıların anlamı şudur:

β_0: Tüm açıklayıcı değişkenler sıfıra eşitken ($X_{1i} = X_{2i} = \ldots = X_{ki} = 0$) bağımlı değişkenin aldığı değerdir.

β_1: X_1 dışındaki tüm açıklayıcı değişkenler sabitken, X_1'deki bir birimlik değişmenin bağımlı değişkende ortaya çıkardığı değişmedir. $E(Y_i|X_{1i}, X_{2i}, \ldots, X_{ki})$'nin X_{1i}'ye göre türevine eşittir.

Herhangi bir β_j: X_j dışındaki tüm açıklayıcı değişkenler sabitken, X_j'deki bir birimlik değişmenin bağımlı değişkende ortaya çıkardığı değişmedir.

Örnek 4.1: $W_i = \beta_0 + \beta_1 E_i + u_i$ modelinde W_i herhangi bir kişinin aylık ücretini ve E_i eğitim yılı sayısıdır. Bu modelin tahmini sonucu $\widehat{W}_i = 1200 + 300E_i$ bulunmuştur. Katsayıların yorumu şu şekildedir:

β_0: Kişinin eğitim yılı sayısı 0 iken yani hiç eğitimi olmayan bir kişinin aylık ücreti 1,200TL'dir.

β_1: Kişinin eğitim yılı bir yıl arttığında ücreti 300TL artar.

Örnek 4.2: $C_t = \beta_0 + \beta_1 Y_t + \beta_2 C_{t-1} + u_t$ modelinde C_t Türkiye'nin herhangi bir t yılındaki tüketim harcamalarını (milyon TL), Y_t ulusal gelirini (milyon TL) ve C_{t-1} bir önceki yılın tüketim harcamalarını (milyon TL) gösterir. Bu model tahmin edilmiş, $\widehat{C}_t = 75 + 0.63Y_t + 0.32C_t$ bulunmuştur. Katsayıların yorumu şu şekildedir:

β_0: Türkiye'de ulusal gelir ve önceki yılın tüketim harcamaları sıfır iken bu yılın tüketim harcaması (otonom tüketim harcaması) 75 milyon TL'dir.

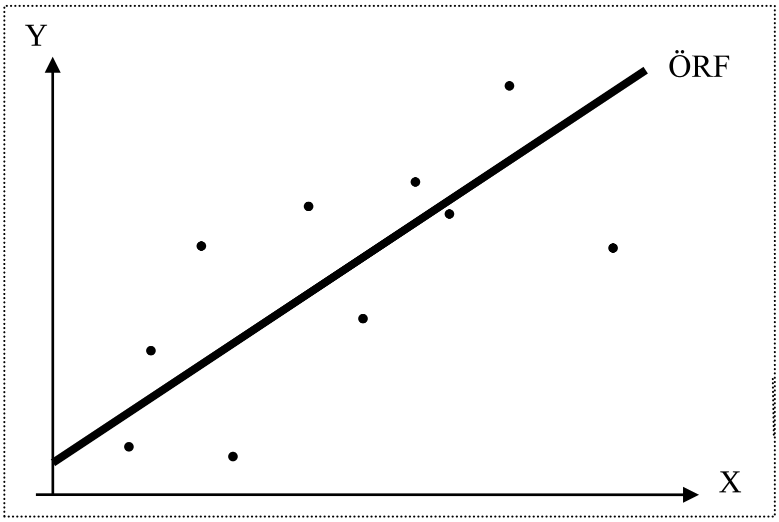
β_1: Bir önceki yılın tüketim harcamasında bir değişme yokken ulusal gelirde 1 milyon TL değişme olduğunda bu yılın tüketim harcamaları aynı yönde 630,000TL değişir. Marjinal tüketim eğilimi 0.63'tür.

β_1: Ulusal gelirde bir değişme yokken bir önceki yılın tüketim harcamasında 1 milyon TL'lik değişme olduğunda bu yılın tüketim harcamalarında 320,000TL'lik artış olur.

4.2. Belirlilik Katsayısı: R^2

Belirlilik katsayısı[1] (R^2) örneklem verileri kullanılarak elde edilen örneklem eğrisinin verilere ne kadar iyi uyduğunu ölçmek amacıyla kullanılan bir ölçüttür. Grafik 4.1 örneklem verilerinin ÖRF fonksiyonu etrafında dağıldığını göstermektedir.

Grafik 4.1: Örneklem Verileri ve Örneklem Eğrisi



En iyi durumda, diğer bir deyişle tam bir uyumun sağlandığı durumda, bütün gözlemler eğri üzerinde olacaktır. Ancak böyle bir duruma rastlama olasılığı çok düşüktür. Genellikle gözlemler eğrinin etrafında dağılacaktır. Gözlemler eğriye ne kadar yakınsa (hata terimleri ne kadar küçükse) o kadar iyi bir uyum sağlanmış olur. Belirlilik katsayısı ise gözlemlerin eğriye ne kadar yakın olduğunu, diğer bir deyişle örneklem regresyon eğrisinin veriye ne kadar iyi uyduğunu gösteren özet bir ölçüdür.

R^2'yi hesaplamak için önce her bir Y değerini tahmin değeri ile hata teriminin toplamı olarak ifade edelim:

$$Y_i = \hat{Y}_i + \hat{u}_i$$

Bunu ortalamadan sapmalar olarak yazarsak

$$(Y_i - \bar{Y}_i) = (\hat{Y}_i - \bar{Y}_i) + \hat{u}_i$$

Karelerini ve toplamlarını alırsak

---

[1] Coefficient of Determination

$$\sum (Y_i-\overline{Y}_i)^2 = \sum (\widehat{Y}_i-\overline{Y}_i)^2 + \sum \hat{u}_i^2$$

Burada $\sum (Y_i-\overline{Y}_i)$ ifadesi Y değerlerinindeki (ortalamalara göre) toplam değişimi gösterdiğinden *Bütün Kareler Toplamı* (BKT), $\sum (\widehat{Y}_i-\overline{\overline{Y}})^2$ ifadesi tahmin edilmiş Y değerlerinindeki (ortalamalara göre) toplam değişimi gösterdiğinden *Açıklanan Kareler Toplamı* (AKT) ve $\hat{u}_i^2$ Y değerlerinin açıklanamayan kısmını gösterdiğinden *Kalıntı Kareler Toplamı* (KKT) olarak adlandırılır. Yeniden ifade etmek gerekirse, BKT Y'deki toplam değişimleri ifade ederken bu değişimlerin yaptığımız tahminin açıklayabildiği kısmı AKT, açıklayamadığı kısmı KKT ile gösterilmiştir.

BKT = AKT + KKT

veya

$$1 = \frac{\text{AKT}}{\text{BKT}} + \frac{\text{KKT}}{\text{BKT}}$$

R^2 değeri, BKT'nın ne kadarının tahminimiz tarafından açıklandığını ölçer, yani AKT/BKT'dir.

$$R^2 = 1 - \frac{\text{KKT}}{\text{BKT}}$$

$$= 1 - \frac{\sum \hat{u}_i^2}{\sum (Y_i-\overline{Y}_i)^2} \tag{4.1}$$

$$= 1 - \frac{\sum \hat{u}_i^2}{\sum Y_i^2 - \frac{(\sum Y_i)^2}{n}}$$

Modelde sabit terim varsa, R^2 değeri 0 ile 1 arasında bir değer alır. Bire ne kadar yakınsa modelin bağımlı değişken Y'deki değişmeleri açıklama gücü o kadar yüksek demektir. Örneğin R^2 0.75 çıkmış ise, model Y'deki değişmelerin yüzde 75'ini açıklamaktadır.

Zaman serileri genellikle trend içerdiğinden R^2 genellikle yüksek çıkmaktadır. Bu nedenle zaman serisi kullanılıyorsa, R^2 0.9 veya üzerinde çıkıyorsa modelin açıklama gücü "yüksek" kabul edilir. Kesit verisinde ise 0.5 veya üzerinde çıkması durumunda açıklama gücü "yüksek" kabul edilir.

Örnek 4.3: Örnek 3.1'de kullanılan verileri ve elde edilen tahmin sonuçlarını kullanarak R^2 değerini hesaplayalım.

Tablo 3.1 ve ve 3.2'deki bilgileri kullanarak R^2 değeri aşağıdaki gibi hesaplanabilir:

$$R^2 = 1 - \frac{\sum \hat{u}_i^2}{\sum Y_i^2 - \frac{(\sum Y_i)^2}{n}} = 1 - \frac{337}{132{,}100 - \frac{(1{,}110)^2}{10}} = 0.962$$

Bulduğumuz sonuca göre model Y'deki değişmelerin yüzde 96.2'sini açıklamaktadır. Kesit verisi kullanıldığı ve elde edilen değer 0.50'nin üzerinde olduğundan modelin açıklama gücü yüksektir diyebiliriz.

4.3. Düzeltilmiş R^2 : $\bar{R}^2$

Modele açıklayıcı değişken eklendikçe R^2 değeri asla azalmaz, genellikle artar. Bu nedenle R^2 değeri, açıklayıcı değişken sayısı aynı olan modellerin açıklama güçlerinin karşılaştırılmasında kullanılmalıdır. Açıklayıcı değişken sayısı farklı olan modelleri karşılaştırmada R^2 değerini kullanmak doğru değildir. Bunun yerine, modele ilgisiz açıkayıcı değişken eklendiğinde bu işlemi cezalandıran alternatif bir istatistik, *Düzeltilmiş R^2* ($\bar{R}^2$) kullanılmalıdır.

$$\bar{R}^2 = 1 - \frac{\hat{u}_i^2/(n-k)}{\sum (Y_i - \bar{Y}_i)^2/(n-1)} = 1 - \frac{\hat{u}_i^2/(n-k)}{(\sum Y_i^2 - \frac{(\sum Y_i)^2}{n})/(n-1)} \qquad (4.2)$$

R^2 formülünü de dikkate alarak 3.15 no'lu denklem R^2 cinsinden de yazılabilir.

$$\bar{R}^2 = 1 - (1 - R^2)\frac{n-1}{n-k} \qquad (4.3)$$

Düzeltilmiş R^2 istatistiğinin özellikleri R^2 ile benzerdir. Farklı olarak modelde sabit terim yer alsa bile düzeltilmiş R^2 eksi değerli olabilir. Ayrıca modele yeni değişken eklendiğinde R^2'nin aksine düzeltilmişi R^2 azalabilir: Yeni değişken eklendiğinde k artacağından (n-1)/(n-k) değeri azalacaktır. Yeni değişkenin Y'yi açıklama gücü düşükse R^2 değeri fazla artmayacağından düzeltilmiş R^2 değeri azalabilir.

Düzeltilmiş R^2 istatistiğinin yorumu R^2 ile aynıdır.

Örnek 4.4: Örnek 3.1'de kullanılan veriler ve tahmin sonuçları ile düzeltilmiş R^2 değerini hesaplayalım.

$$\bar{R}^2 = 1\text{-}(1\text{-}R^2)\frac{n-1}{n-k} = 1\text{-}(1\text{-}0.962)\frac{10-1}{10-2} = 0.957$$

Düzeltilmiş R^2 değerine göre model Y'deki değişmelerin yüzde 95.7'sini açıklamaktadır.