8. Hafta İçsellik

Klasik doğrusal regresyon modelinin temel varsayımlarından biri, X bağımsız değişkeninin değerinin tekrarlanan örneklerde değişmediği diğer bir ifade ile rassal olmadığıdır. X bağımsız değişkeninin rassal olmadığı varsayımının nedenleri:

- En küçük kareler cebirini basitleştirmektedir.
- X rassal olsa bile, en küçük kareler tahmin edicinin özellikleri çok az değiştirilmiş varsayımlar altında hala geçerli olmaktadır.

Klasik doğrusal regresyon modelinin temel varsayımlarından bir diğeri ise, X bağımsız değişkeni(leri) ile rassal hatanın ilişkisiz olduğudur. $E(u|X_i)=0$ ise, X ile u 'nin korelasyonsuz olduğu aşağıdaki gibi gösterilir.

$$Cov(X_i, u_i) = 0$$

1. Rassal X 'ler ile Doğrusal Regresyon

Basit regresyon modelinin temel varsayımlarını aşağıdaki gibi değiştirelim:

1. $Y_i = \beta_0 + \beta_1 X_i + u_i$, anakütlede Y_i ve X_i arasındaki ilişki doğrusal olarak tanımlanmaktadır, burada β_0 ve β_1 bilinmeyen (sabit) ana kütle parametreleri, u_i ise gözlemlenemeyen rassal hata terimidir.
2. Veri çiftleri (X_i, Y_i) $i = 1, 2, \ldots, n$ için **rassal örnekleme** ile elde edilmektedir. Aynı ana kütleden gelen veri çiftleri, diğer veri çiftinden bağımsızdır. Kısaca bağımsız ve aynı dağılımlıdır.
3. $E(u|X)=0$, X herhangi bir değeri için ana kütle hata terimi u 'nin beklenen değeri (koşullu ortalaması) sıfırdır.
4. Örneklemde, X en azından iki farklı değer almalıdır, $Var(X)>0$.
5. $Var(u|X)=\sigma^2$, hata teriminin koşullu varyansı bir sabit σ^2 'dir.
6. Hata terimi normal dağılmaktadır, $u_i \sim N(0,\sigma^2)$.

Yukarıdaki varsayımlardan 2. varsayım hem Y 'nin hem de X 'in bir örnekleme süreci ile elde edildiğini ve böylece rassal olduğunu belirtmektedir. Ayrıca, çiftlerin bağımsız olduğu varsayımı,

$$Cov(u_i, u_j) = Cov(Y_i, Y_j) = 0 \qquad i \neq j$$

varsayımının da geçerli olduğu anlamına gelmektedir.

$E(u|X)=0$ varsayımı

(i) Önemli bir değişkenin dışlanmış olmadığı yani model dışında kalmadığı,

(ii) Fonksiyonel biçimin doğru olduğu

(iii) Hata terimi u 'nin X ile korelasyonlu olmadığı, $Cov(X_i, u_i)=0$.

anlamına gelmektedir.

Hata terimi ile korelasyonlu olmayan bağımsız değişkenler **dışsal değişken** olarak adlandırılmaktadır. Dışssal değişken "bir sistemin dışında belirlenen" anlamına gelmektedir. Örneğin, bir arz ve talep modelinde, kasırgalar, depremler ve salgın hastalıklar gibi olaylar yüzünden arzdaki değişmeler "dışsaldır".

X ve u korelasyonlu $Cov(X_i, u_i) \neq 0$ ise, bunun sonucu olarak $E(u|X) \neq 0$ 'dir. **Hata terimi ile korelasyonlu** olan bağımsız değişkenler ise **içsel değişken** olarak adlandırılır. Bu "bir sistem içinde belirlenen" anlamına gelmektedir. Örneğin, bir arz ve talep modelinde denge fiyatı ve miktarı içseldir.

1.1. En Küçük Kareler Tahmin Edicinin Küçük Örneklem Özellikleri

Gauss-Markov teoremine göre, klasik doğrusal regresyon modelinin temel varsayımları geçerli ve X 'ler sabit ise, en küçük kareler tahmin edicileri en iyi (tesirli, en küçük varyanslı) doğrusal ve eğilimsiz (sapmasız) tahmin edicilerdir (BLUE). Sonuç, örneklem büyüklüğüne bağlı değildir, (sonlu örneklem veya küçük örneklem). Örneklem büyüklüğü n=20, 50 veya 10,000 olmasına bağlı kalmaksızın, söz konusu durum her örneklemde geçerli olmaktadır.

Klasik doğrusal regresyon modelinin temel varsayımları geçerli ancak X 'in rassal olduğu durumda en küçük kareler tahmin edicinin sonlu örneklem özellikleri aşağıdaki gibidir.

1. En küçük kareler tahmin edicisi, regresyon parametrelerinin en iyi, doğrusal, eğilimsiz tahmin edicisidir ve σ^2'nin tahmin edicisi eğilimsizdir, $E(\hat{\sigma}^2)=\sigma^2$.
2. En küçük kareler tahmin edicilerinin dağılımları normaldir. Aralık tahmini ve hipotez testi prosedürleri geçerlidir.

Buna göre, X rassal ise, verinin rassal örneklem ile elde edildiği ve diğer temel varsayımlar geçerli olduğu sürece, regresyon yönteminde değişime gerekli değildir.

1.2. En Küçük Kareler Tahmin Edicinin Büyük Örneklem Özellikleri

Çoklu regresyon modelinde, varsayımlar temel varsayımlar sağlanırsa, en küçük kareler tahmincilerinin en iyi, doğrusal, eğilimsiz tahminciler (BLUE) olması özellikleri örnek büyüklüğü n'e bağlı değildir ve bu özellik "sonlu örnek" özellikleri olarak adlandırılır. $n > k$ olduğu tüm durumlarda geçerlidir.

Hatırlatma (Temel varsayımlar)

Basit regresyon modeli için:

1. X in her bir değeri için, Y 'nin değeri,

$$Y_i = \beta_0 + \beta_1 X_i + u_i$$

2. Rassal hata u 'nin beklenen değeri,

$$E(u_i) = 0$$

Bu varsayım, $E(Y) = \beta_0 + \beta_1 X_i$ varsayıma eşdeğerdir:

3. Rassal hata u 'nin varyansı,

$$Var(u_i) = \sigma^2 = Var(Y)$$

4. Rassal hatalar u_i ve u_j herhangi bir çifti arasındaki kovaryans,

$$Cov(u_i, u_j) = Cov(Y_i, Y_j) = 0 \qquad i \neq j$$

Bu varsayımın daha güçlü biçimi, rassal hatalar u 'nin istatistiksel olarak bağımsız olması halidir, -ki bu durumda bağımlı değişken Y nin değerleri de istatiksel olarak bağımsızdır.

5. X değişkeni rassal değildir ve en azında iki farklı değer almak zorundadır, $Var(X) > 0$.

Çok değişkenli regresyon modeli için:

1. $Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + +\beta_{(k-1)i} X_{(k-1)i} + u_i \qquad i = 1, 2, \ldots, n$

2. $E(Y_i) = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + +\beta_{(k-1)i} X_{(k-1)i} \Leftrightarrow E(u_i) = 0$

3. $Var(u_i) = \sigma^2 = Var(Y)$

4. $Cov(u_i, u_j) = Cov(Y_i, Y_j) = 0 \qquad i \neq j$

5. X değişkenleri rassal değildir ve diğer bağımsız değişkenlerin tam doğrusal bir fonksiyonu değildir.

Örneklerin son derece büyük olduğu düşünüldüğünde, en küçük kareler tahmincisinin ek özellikleri, $n \longrightarrow \infty$ ile gösterilen **asimptotik özellikler** olarak adlandırılır. Bunlar

1. Tutarlılık

Ekonometrik tahminci seçilirken, gerçek olan fakat bilinmeyen ana kütle parametresine yüksek olasılıkla yakın bir tahmini elde etmek temel hedeftir. Basit regresyon modeli

$$Y_i = \beta_0 + \beta_1 X_i + u_i \qquad i = 1, 2, \ldots, n$$

gözönüne alındığında, β_1 'in gerçek değerinin "epsilon" sınırları içinde olduğu varsayılır ve ana kütle parametresi β_1'ye "yakın" bir tahmin elde etme olasılığı aşağıdaki gibidir,

$$P(\beta_1 - \varepsilon \leq \hat{\beta}_1 \leq \beta_1 + \varepsilon)$$

Örnek büyüklüğü $n \longrightarrow \infty$ giderken, bu olasılık 1'e yaklaşıyorsa bu tahminci **tutarlı**dır. veya limit kavramı kullanılarak $\hat{\beta}_1$ tahmincisi aşağıdaki koşulda **tutarlıdır**,

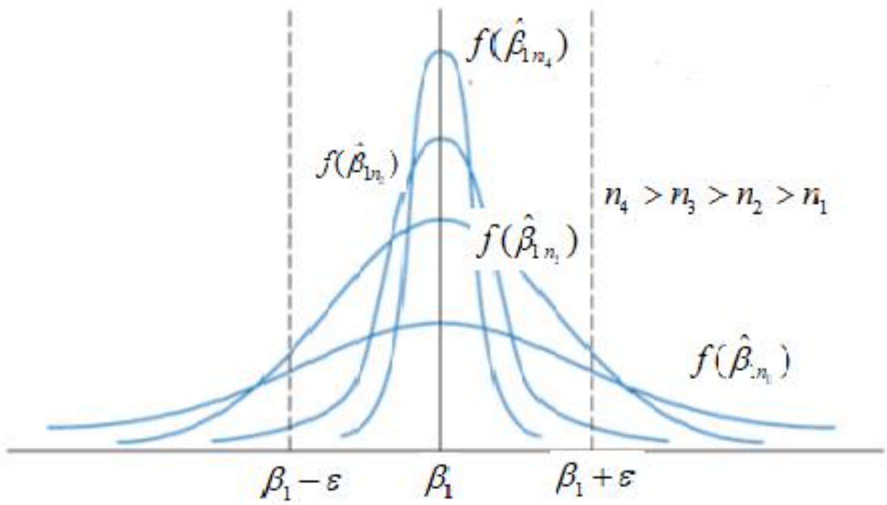
$$\lim_{n \to \infty} P(\beta_1 - \varepsilon \leq \hat{\beta}_1 \leq \beta_1 + \varepsilon) = 1$$

Aşağıdaki şekilde, n örnek büyüklüğü ve $n_4 > n_3 > n_2 > n_1$ olmak üzere, en küçük kareler tahmincisi $\hat{\beta}_1$ için olasılık dağılım fonksiyonları $f(\hat{\beta}_{1,n_i})$ 'dir. Örnek büyüklüğü arttıkça, olasılık dağılım fonksiyonu daha dar olacaktır. Temel varsayımlar sağlandığında en küçük kareler tahmincisi eğilimsizdi, $E(\hat{\beta}_1) = \beta_1$. Bu özellik, tüm örnek büyüklükleri için geçerlidir. Örnek büyüklüğü değiştikçe, olasılık dağılım fonksiyonunun merkezi, β_1'de kalır.

Basit regresyon modelinde ($Y_i = \beta_0 + \beta_1 X_i + u_i$) en küçük kareler tahmincisi $\hat{\beta}_1$ 'nin varyansı $Var(\hat{\beta}_1) = \sigma^2 \big/ \sum x_i^2$ ve çok değişkenli regresyon modelinde ($Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + u_i$ ise $Var(\hat{\beta}_1) = \sigma^2 \big/ (1 - r_{12}^2) \sum x_i^2$ 'dir. Her iki durumda da örnek büyüklüğü (n) arttıkça, tahminci $\hat{\beta}_1$ 'nin varyansının daha küçük olacaktır. n arttıkça, olasılık dağılım fonksiyonlarının merkezi, $E(\hat{\beta}_1) = \beta_1$'de sabitlenmiş olarak kalır ve varyans küçülür.

$\hat{\beta}_1$'nin $\beta_1 - \varepsilon \leq \hat{\beta}_1 \leq \beta_1 + \varepsilon$ aralığına düşme olasılığı, buu limitler arasındaki olasılık dağılım fonksiyonunun altındaki alandır. Örnek büyüklüğü arttıkça, $\hat{\beta}_1$'nin limitler içine düşme olasılığı

1'e yaklaşır. Büyük örneklerde en küçük kareler tahmincisi, yüksek olasılıkla gerçek parametre değerine yakın bir tahmin verecektir.



Tutarlılık özelliği sonlu örneklerde eğilimli olsa bile, pek çok tahminciye uygulanır. Eğilim miktarı aşağıdaki gibidir:

$$\text{eğilim}(\hat{\beta}_1) = E(\hat{\beta}_1) - \beta_1 = \frac{1}{n}$$

$\hat{\beta}_1$ tahmincisi için, $n \longrightarrow \infty$ gittikçe, eğilim sıfıra yakınsar. Yani,

$$\lim_{n\to\infty} \text{eğilim}(\hat{\beta}_1) = \lim_{n\to\infty}\left[E(\hat{\beta}_1) - \beta_1\right] = 0$$

Bu durumda, tahminci **asimptotik olarak eğilimsizdir**. Bir tahminci için tutarlılık, tahmincinin ya eğilimsiz ya da asimptotik olarak eğilimsiz olduğunu, yani $n \longrightarrow \infty$ gittikçe, varyansın sıfıra yakınsadığı anlamına gelmektedir.

$$\lim_{n\to\infty} Var(\hat{\beta}_1) = 0$$

Tutarlı bir tahmincinin olasılık yoğunluk fonksiyonu, parametrenin gerçek değeri etrafında yoğunlaştığı ve tahminci $\hat{\beta}_1$'nin, gerçek parametre β_1'ye yaklaşma olasılığı bire yaklaştığı için, tahminci $\hat{\beta}_1$ 'nin, β_1'ye "olasılıkta yakınsadığı" söylenebilir.

2. Asimptotik Normallik

Yukarıdaki temel varsayımları sağlayan model için ekonometrik tahminciler büyük örneklerde yaklaşık bir normal dağılıma sahiptir. $n \longrightarrow \infty$ gittikçe, standardize edilmiş tahmincinin olasılık yoğunluk fonksiyonu aşağıdaki standart normal dağılıma yaklaşan bir dağılıma sahiptir,

$$\frac{\hat{\beta}_i - \beta_i}{\sqrt{Var(\hat{\beta}_i)}} \overset{a}{\sim} N(0,1)$$

Bu durumda, tahmincinin asimptotik olarak normal dağılır ve aşağıdaki gibi gösterilir.

$$\hat{\beta}_i \overset{a}{\sim} N(\beta_i, Var(\hat{\beta}_i))$$

Bu sonuç, Merkezi Limit Teorimi ile benzerdir.

Bu sonucun önemi, $u_i \sim N(0,\sigma^2)$ varsayımı sağlanmasa bile, örnek yeterince büyük ise *t*-testleri, *F*-testleri ve genel aralık tahmini ve öngörü aralığının uygulanabilir olmasıdır.”Yeterince büyük örnek sayısı” ise her bir uygulama için verilerin ve hata teriminin yapısına bağlı olarak değişiklik gösterir. Daha karmaşık modelin, yaklaşık normalliği sağlaması için daha büyük bir örnek gerekecektir.

En küçük kareler tahmin edicinin bir “büyük örneklem” analizinin amaçları için, yukarıdaki varsayımlardan 3. Varsayım olan $E(u|X)=0$’ı , Varsayım 3* $E(u)=0$ ve $Cov(X,u)=0$ ile değiştiriyoruz. 3. varsayım doğru ise, $E(u|X)=0 \Rightarrow Cov(X,u)=0$ ve $E(u|X)=E(u)=0$ dır. Şöyle ki;

Regresyon model $Y_i = \beta_0 + \beta_1 X_i + u_i$’de, Y’nin koşullu ortalaması $E(Y|X)=\beta_0+\beta_1 X_i$ ve eş değer olarak $E(u|X)=0$’dır. X üzerine koşullu hata teriminin beklenen değeri sıfırdır. *X*’lerin rassal olduğu durum için

$$E(u) = E_X\left[E(u|X)\right] = E_X\left[0\right] = 0$$

$$E(Xu) = E_X\left[XE(u|X)\right] = E_X\left[X \times 0\right] = 0$$

$$Cov(X,u) = E_x\left[(X-\mu)E(u|X)\right] = E_x\left[(X-\mu)0\right] = 0$$

Böylece, $E(u|X)=0$ ise bu $E(u)=0$, $E(Xu)=0$ ve $Cov(X,u)=0$ olduğunu izlemektedir, $E(u) \neq 0$ ise’ $Cov(X,u) \neq 0$’dır.

Ekonometrik modelin tahmin süreci açısından, rassal bağımsız değişken X ile hata terimi u korelasyonlu olduğu durum –yani $Cov(X,u)=0$ varsayımının ihlal edilmesi- önem kazanmaktadır.

Varsayım 3* gerçekte Varsayım .3'den daha zayıf bir varsayımdır. Varsayım 3* altında en küçük kareler tahmin edicinin sapmasız olduğu veya diğer sonlu örneklem özelliklerinin her hangi birinin geçerli olduğu gösterilemez.

Varsayım 1, 2, 3*,.4 ve 5 altında, en küçük kareler tahmin ediciler

1. Tutarlıdır. Yani, onlar $n \to \infty$ iken olasılıkta gerçek parametre değerlerine yaklaşmaktadır.
2. Hatalar normal olarak dağılsa da dağılmasa da, tahmin edilen parametreler büyük örneklemlerde yaklaşık normal dağılımlara sahiptirler. Ayrıca, örneklem büyük ise olağan aralık tahminlerimiz ve test istatistiklerimiz geçerlidir.
3. Varsayım 3* doğru değil ise ve özelliklede $KOV(X,u)\neq 0$ ise böylece X ve u korelasyonludur, bu durumda en küçük kareler tahmin ediciler tutarsızdır. $\hat{\beta}_i$'ler gerçek parametre (β_i) değerlerine çok büyük örneklemlerde bile yaklaşmazlar. Ayrıca, hipotez testi ve aralık tahmin prosedürleri geçerli değildir.

X'in rassal olduğu durumda, en küçük kareler tahmininin uygun olup olmadığına karar verirken önemli bir faktör X ve u arasındaki ilişkidir. Hata terimi u, X (veya çoklu regresyon modelinde her hangi bir X) ile korelasyonlu ise, en küçük kareler tahmin edici başarısız olmaktadır.

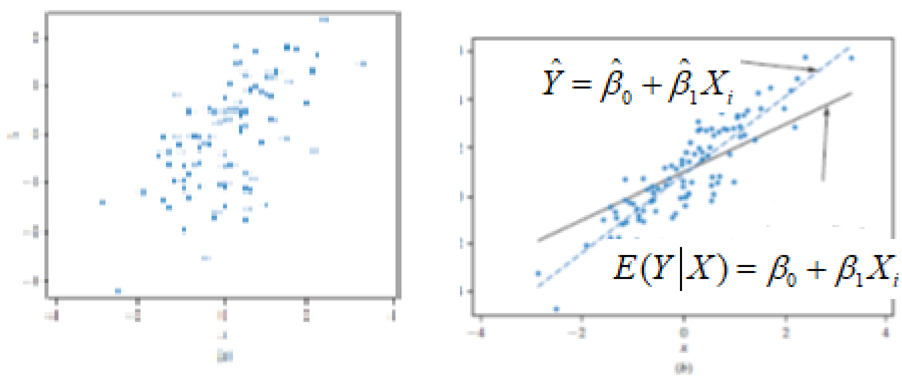
1.3. X ve u arasındaki korelasyonlu ise neden en küçük kareler tahmin edicisinin başarısız olmaktadır.

Regresyon modeli veri üretme sürecinde, gözlemlenen sonuç Y_i' yi elde etmek için sistematik kısım $E(Y|X)=\beta_0+\beta_1 X_i$ 'e rassal hata u eklemektedir.

$$Y_i = E(Y|X)+u_i \qquad Y_i=\beta_0+\beta_1 X_i+u_i$$

Aşağıdaki şekil (a)'da X ve u değerleri pozitif olarak korelasyonludur. Şekil 1 (b)'de düz çizgi ile gösterilen pozitif-eğimli ana kütle regresyon fonksiyonu $E(Y|X)=\beta_0+\beta_1 X_i$ dur. X'in her bir değeri için, Şekil (b)'de noktalar olarak gösterilen, Gözlemlene Y değerleri ile

$E(Y|X) = \beta_0 + \beta_1 X_i$ arasındaki fark rassal hata u kadardır, $Y_i = \beta_0 + \beta_1 X_i + u_i$. Görüldüğü üzere, anakütle regresyon doğrusu gözlemlenen verilerin ortasından geçmemektedir. Bu durum X ve u arasındaki korelasyondan kaynaklanmaktadır. Daha büyük X değerleri için Y değerleri pozitif hatalar u 'ye sahiptir. Daha küçük X değerleri için Y değerleri negatif hatalar u 'ye sahiptir.



Şekil (a) Korelasyonlu X ve u (b) Veri noktaları, gerçek ve tahmini regresyon fonksiyonları.

Şekil (b)'de kesikli çizgi olarak gösterilen, en küçük kareler tahmini verinin ortasından geçen uyumlu bir çizgi ile sonuçlanmaktadır. EKK yöntemi ile tahmin edilen doğrunun eğimi $\hat{\beta}_1$, ana kütle regresyon doğrusunun eğiminden $\beta_1 > 0$ daha büyüktür. Y 'deki değişimin iki kaynağa dayanmaktadır. X 'teki değişimler ve u 'deki değişimler ve bu değişimler pozitif bir korelasyona sahiptir. X ve u 'daki değişimlerin Y üzerinedeki etkisi

$$\Delta Y = \beta_1 \Delta X + \Delta u$$

(+) (+) (+)

X ve u pozitif korelasyonlu ve $\beta_1 > 0$ ise, X ve u değerlerindeki artışlar Y yi arttırmak için birleşmektedir. En küçük kareler tahmin sürecinde, Y deki bütün değişim (artış) X 'teki değişimin (artışın) etkisine dayandırılmaktadır ve böylece en küçük kareler tahmin edici β_1 'yi aşırı tahmin edecektir.

X ve u arasında korelasyon varsa en küçük kareler tahmin edicisi **eğilimlidir** ve bu eğilim örneklem büyüklüğüne bağlı değildir, örnek sayısını arttırmak eğilimsiz tahminciler elde etmeyi sağlamayacaktır. Sonuç olarak X ve u arasında korelasyon varsa, en küçük kareler tahmin edici **tutarsızdır**.

1.4. X ve u'nin Korelasyonlu Olduğundaki Durumlar

Bağımsız bir değişken ve hata terimi korelasyonlu ise, bağımsız değişken içsel olarak nitelendirilmektedir. Bu nitelendirme eşanlı denklem modellerinden gelmektedir ve "sistem içinde belirlenen" anlamına gelmektedir. Bağımsız değişken, rassal hata ile korelasyonlu olduğu durumda "içsellik problemi"nin varlığından söz edilir. Sebepleri:

1. **Ölçüm Hatası**

Bağımsız değişken hatalı ölçüldüğü durumda, rassal hata ile korelasyonludur ve en küçük kareler tahmin edici tutarsızdır. Bir bireyin kişisel tasarrufu, tüketimi gibi, bireyin "sürekli" veya uzun dönemde gelirinin fonksiyonu olduğunu varsayalım. Y kişinin yıllık tasarrufları, X^* ise sürekli yıllık geliri olsun. Bu ilişkiyi temsil eden basit bir regresyon modeli:

$$Y = \beta_0 + \beta_1 X^* + v$$

X^* ile gösterilen sürekli gelir değişkenini gözlemlemek, imkânsız değilse, zordur.Regresyonun amaçları için, sürekli gelir yerine cari geliri (X) kullandığımızı varsayalım. Cari gelir sürekli gelirin bir ölçümüdür, fakat bu sürekli geliri tam olarak ölçmemektedir. Buna bazen **vekil bir değişken** denilmektedir.

$$X = X^* + u$$

burada u, 0 ortalama ve varyans σ_u^2 ile rassal hatadır. Gözlemlenen cari gelirin sadece sürekli gelire yaklaştığını ve sonuç olarak hata ile sürekli geliri ölçmüş olduğumuzu kabul ediyoruz. Ayrıca, u'nun v'den bağımsız olduğunu ve korelasyonsuz olduğu varsayılmakatdır. Regresyonda X^*'ın yerinde X 'i kullandığımız zaman, bunu yer değiştirme ile gerçekleştirmekteyiz. Yani,

$$Y = \beta_0 + \beta_1 X^* + v$$

$$Y = \beta_0 + \beta_1 (X - u) + v$$

$$Y = \beta_0 + \beta_1 X + (v - \beta_1 u)$$

$$Y = \beta_0 + \beta_1 X + e$$

$Y = \beta_0 + \beta_1 X + e$ deki bağımsız değişken X , $X = X^* + u$ ölçüm hatasının varsayımından dolayı rassaldır.

$Y = \beta_0 + \beta_1 X + e$ modelini En küçük kareler ile tahmin etmek için, X 'in rassal hata e ile korelasyonlu olup olmadığını tespit edilmelidir. $E(e) = 0$ varsayımını kullanarak bu iki rassal değişken arasındaki kovaryans

$$Cov(X,e) = E(Xe) = E(X^* + u)(v - \beta_1 u)$$

$$= E(-\beta_1 u^2) = -\beta_1 E(u^2) = -\beta_1 \sigma^2 \neq 0$$

sıfırdan farklıdır. En küçük kareler tahmin edici $\hat{\beta}_1$ açıklayıcı değişken ve hata terimi arasındaki korelasyondan dolayı β_1 'nin *tutarsız* bir tahmin edicisidir. Sonuç olarak, $\hat{\beta}_1$ büyük örneklemlerde β_1'ye yaklaşmamaktadır. Ayrıca, büyük veya küçük örneklemlerde $\hat{\beta}_1$, ortalama β_1 ve varyans $Var(\beta_1) = \sigma^2 / \sum x_i^2$ ile yaklaşık olarak normal dağılıma uygun değildir. En küçük kareler bu şekilde başarısız olduğu zaman, moment yöntemine dayalı tahmin ediciler kullanılmaktadır.

2. Eşanlı Denklemler Sapması

Bağımsız değişken ile rassal hata teriminin korelasyonlu olduğunda başka bir durum eşanlı denklemler modellerinde ortaya çıkmaktadır. Rekabetçi piyasada, malın fiyatı ve miktarı arz ve talep tarafından ortaklaşa belirlenmektedir. Böylece $P =$ denge fiyatı ve $Q =$ denge miktarı ise P ve Q'nun içsel olduğunu söyleyebiliriz, çünkü onlar arz eğrisi için bir denklem ve talep eğrisi için diğer denklem olan iki denklemli eşanlı bir sistem içinde ortaklaşa belirlenmektedir.

$$Q = \beta_0 + \beta_1 P + u$$

Bir yandan fiyattaki değişim, arz ve talep miktarlarını etkilerken diğer yandan arz ve talep edilen miktarlardaki değişimler de fiyatta değişimlere etkilemektedir. P ve Q arasında bir karşılıklı ilişki (feedback) ilişkisi vardır. Fiyat ve miktarın ortaklaşa veya eşanlı olarak belirlenmesinden dolayı sonuçlanan bu karşılıklı etkilşimden dolayı, $Cov(P,u) \neq 0$ olacaktır. En küçük kareler tahmin prosedürü $Q = \beta_0 + \beta_1 P + u$ uygulanırsa içsellik probleminden dolayı başarısız olacaktır ve meydana gelen sapmaya (ve tutarsızlığa) eşanlı denklemler sapması denilmektedir.

3. İhmal Edilen Değişkenler

Model dışında kalan değişken (dışlanmış değişken) dâhil edilen, modelde yer alan bağımsız bir değişken ile korelasyonlu olduğu durumda regresyon hatası içsel olan açıklayıcı değişken ile korelasyonlu olacaktır. Klasik bir örnek çalışma ekonomisindendir. Bir kişinin ücreti kısmen onun eğitim düzeyi ile belirlenmektedir. X_1 eğitim yılını ve X_2 deneyim yılını göstermek üzere ücret (Y) denklemi aşağıdaki gibidir.

$$\ln Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_2^2 + u \qquad \text{log-doğrusal regresyon modeli}$$

Ücretleri belirleyen işgücü piyasası koşulları, bölge ve sendika üyeliği gibi çeşitli faktörler modele dahil edilmemiş ve dışlanmıştır. Ayrıca , çalışan kişinin yeteneği ölçen bir değişken de modelde yer almamış dışlanmıştır. Bir kişinin yeteneğinin (ve çalışkanlığının) işin kalitesini ve ücretleri etkileyebilir olması mantıklıdır. Genellikle ölçüme sahip olmadığımız için bu değişkenler hata terimi u'nin bileşenleridir. Problem yeteneğin sadece ücretleri etkileyebildiği değildir, aynı zamanda hata terimi u ve eğitim değişkeni X_1 arasında pozitif bir korelasyona neden olan, daha yetenekli bireylerin de okulda daha fazla yıllar harcayabilmesidir, bu nedenle $Cov(X_1, u) > 0$ dır. Eğer bu doğru ise, o zaman eğitimin bir diğer yıl getirilerinin en küçük kareler tahmin edicinin pozitif eğilimli, $E(\hat{\beta}_1) \neq \beta_1$ ve tutarsız olacağını bekleyebiliriz, bunun anlamı sapmanın çok büyük örneklemlerde bile kaybolmayacağıdır.

Ücret Denkleminin En Küçük Kareler Tahmini

Bu bölümdeki uygulamalar ücretler ile eğitim yılı ve işyeri deneyimi arasındaki ilişkiyi kuran regresyon modeli çerçevesinde gerçekleştirilecektir. Modelde dışlanmış (ihmal edilmiş) değişken "zeka", regresyon hatasındadır. Kişinin zekası eğitim aldığı yıllar ile muhtemelen pozitif olarak korelasyonludur. Ücretin bağımlı değişken eğitim yıllarının bağımsız değişken olduğu modelde, en küçük kareler tahmin edici ile ücretlerdeki artışlar eğitimdeki artışa dayandırılmaktadır. Ücretlerdeki artışın birazı da yüksek zekâya bağlı olduğu için eğitimin etkisi gerçek değerinden daha büyük çıkacaktır.

Örneklemde işgücüne katılan $n = 428$ evli kadına ait verilerin kullanıldığı ücret modelinin en küçük kareler tahminleri ve standart hataları aşağıda verilmiştir.

$$\begin{aligned} &\ln(Wage) = -0{,}522 + 0{,}1075 Educ + 0{,}0416 Exper - 0{,}0008 Exper^2 \\ &(Se) \qquad (0{,}1986) \quad (0{,}0141) \qquad (0{,}0132) \qquad\qquad (0{,}0004) \end{aligned}$$

Diğer değişkenler sabit tutulduğunda, eğitimin ek bir yılının ücretleri yaklaşık olarak %10.75 arttırdığını tahmin edilmiştir. Eğer yetenek ücretler üzerinde pozitif bir etkiye sahip ise, bu durumda bu tahmin, yetenek katkısı eğitim değişkenine dayandırıldığı için abartılmaktadır.

2. Momentler Yöntemine Dayalı Tahmin Ediciler

Basit doğrusal regresyon modelinde $Y_i = \beta_0 + \beta_1 X_i + u_i$, X rassal ve $Cov(X_i, u_i) = E(X_i, u_i) \neq 0$ ise, en küçük kareler tahmin edicileri eğilimli ve tutarsızdır. Bu durumda EKK yöntemine alternatif tahmin yöntemi "Momentler Yöntemi" dir.

- Doğrusal regresyom modelinin temel varsayımlarının hepsi geçerli ise, momentler yöntemi bizi en küçük kareler tahmin edicisine götürmektedir.
- Eğer X rassal ve hata terimi ile korelasyonlu ise, bu durumda momentler yöntemi büyük örneklemlerde kullanılacak olan araç değişkenler (AD) tahmini veya iki-aşamalı en küçük kareler (2AEKK) tahmini bir alternatife götürmektedir.

2.1. Anakütle Ortalaması ve Varyansının Momentler Yöntemi Tahmini

Rassal değişken Y'nin k'ncı momenti, k'ncı kuvvete yükseltilen rassal değişkenin beklenen değeridir. Yani,

$E(Y^k) = \mu_k$ Y'nin k'ncı momenti

k'ncı anakütle momenti, örneklem (n büyüklüğünün) verilerinden tutarlı olarak tahmin edilebilir

$E(Y^k) = \hat{\mu}_k$ Y'nin k'ncı örneklem momenti $= \sum Y_i^k / n$

Momentler yöntemi tahmin prosedürü, m sayıda bilinmeyen parametreyi tahmin etmek için m sayıda anakütle momentini m sayıda örneklem momentine eşitlemeye dayanmaktadır. Y rassal değişkeninin ortalaması ve varyansı:

$$E(Y) = \mu$$

$$Var(Y) = \sigma^2 = E(Y-\mu)^2 = E(Y^2) - \mu^2$$

Yukarıdaki iki anakütle parametresi μ ve σ^2 tahmin etmek için, bu iki anakütle momentinin iki örneklem momentine eşitlenmesi gerekir. Y'nin ilk iki anakütle ve örneklem momenti aşağıdaki gibidir.

Anakütle momentleri Örneklem momentleri

$$E(Y) = \mu_1 = \mu \qquad \hat{\mu} = \sum Y_i / n$$
$$E(Y^2) = \mu_2 \qquad \hat{\mu}_2 = \sum Y_i^2 / n$$

Anakütle ortalamasının (1.moment) tahmini için, birinci örneklem momenti birinci anakütle momentine eşitlenir.

$$\hat{\mu} = \sum Y_i / n = \bar{Y}$$

Anakütle varyansının (2.moment) tahmini için

$$\hat{\sigma}^2 = E(Y^2) - \mu^2 \qquad \text{burada} \qquad E(Y^2) = \hat{\mu}_2 = \sum Y_i^2 / n \qquad \text{ve} \qquad \mu^2 = \hat{\mu} = \sum Y_i / n$$

$$\hat{\sigma}^2 = \hat{\mu}_2 - \hat{\mu}^2 = \frac{\sum Y_i^2}{n} - \bar{Y}^2 = \frac{\sum Y_i^2 - n\bar{Y}^2}{n} = \frac{\sum (Y_i - \bar{Y})^2}{n}$$

Momentler yöntemi anakütle ortalamasının tahmin edicisi, örneklem ortalamasıdır. Momentler yöntemi varyans tahmin edicisi ise paydasında $n-1$'yerine n'e sahiptir. Ancak bu durum büyük örneklemlerde önemli bir fark yaratmaz. Momentler yöntemi tahmin edicileri **tutarlıdır** ve büyük örneklemlerde gerçek parametre değerlerine yaklaşmaktadır, fakat onların herhangi bir anlamda "en iyisi" olduğunun garantisi yoktur.

2.2. Basit Doğrusal Regresyon Modelinde Momentler Yöntemi Tahmini

Bir "moment'' tanımı daha genel durumlar için genişletilebilmektedir. Doğrusal regresyon modeli $Y_i = \beta_0 + \beta_1 X_i + u_i$ için genellikle

$$E(u) = 0 \Rightarrow E(Y_i - \beta_0 - \beta_1 X_i) = 0 \qquad (1)$$

olduğu varsayılmaktadır. Ayrıca, X sabit ve rassal hata u ile korelasyonlu değil ise,

$$E(X,u) = 0 \Rightarrow E\left[X(Y_i - \beta_0 - \beta_1 X_i)\right] = 0 \qquad (2)$$

dır. Yukarıdaki Denklem 1 ve Denklem 2 moment koşullarıdır. İki anakütle momentine karşılık gelen örneklem momentleri ile yer değiştirilirse, β_0 ve β_1 için momentler yöntemi tahmin edicilerini tanımlayan iki bilinmeyenli iki denklem elde edilecektir.

$$\frac{1}{n}\sum (Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i) = 0$$

$$\frac{1}{n}\sum X(Y_i-\hat{\beta}_0-\hat{\beta}_1X_i)=0$$

Bu iki denklem en küçük kareler “normal” denklemlerine eşdeğerdir ve çözümleri en küçük kareler tahmin edicilerini vermektedir.

$$\hat{\beta}_1=\frac{\sum(Y_i-\bar{Y})(X_i-\bar{X})}{\sum(X_i-\bar{X})^2}=\frac{\sum y_i x_i}{\sum x_i^2}$$

$$\hat{\beta}_0=\bar{Y}-\hat{\beta}_1\bar{X}$$

Böylece temel varsayımlar altında, basit doğrusal regresyon modeli için momentler yöntemi tahmin ediciler, en küçük kareler tahmin edicileri ile aynıdır.

2.3. Basit Doğrusal Regresyon Modelinde Araç Değişkenler Tahmini

X rassal ve rassal hata u ile korelasyonlu olduğu durumda en küçük kareler yöntemi için problemler ortaya çıkmaktadır, bu nedeni ise $E(Xu)\neq 0$’dır. Bu durum $E(X,u)=0\Rightarrow E\left[X(Y_i-\beta_0-\beta_1X_i)\right]=0$ moment koşulunu geçersiz kılmaktadır.

Bu durumda, aşağıdaki özelliklere sahip başka bir Z değişkeni varsa:

1. Z değişkeni Y üzerinde direk bir etkiye sahip değildir ve böylece bağımsız bir değişken olarak modelin sağ tarafında yer almamaktadır.
2. Z değişkeni, rassal hata terimi u ile korelasyonlu değildir, kısaca dışsaldır.
3. Z değişkeni, içsel bağımsız değişken olan X ile güçlü (veya en azından zayıf değil) korelasyonludur.

Bu özelliklere sahip Z değişkeni **araç (alet) değişken** olarak adlandırılır. Z değişkeni, Y bağımlı değişkeni üzerinde direk bir etkiye sahip olmamasına ragmen, X bağımsız değişkeni ile Y bağımlı değişkeni arasındaki ilişkinin tahmin edilmesine imkan vermektedir.

Böyle bir Z değişkeninin varlığı durumunda, moment koşulu aşağıdaki gibi yazılır.

$$E(u)=0\Rightarrow E(Y_i-\beta_0-\beta_1X_i)=0$$

$$E(\mathrm{Z},u)=0\Rightarrow E\left[Z(Y_i-\beta_0-\beta_1X_i)\right]=0$$

Daha sonra β_0 ve β_1 tahminlerini elde etmek için, yukarıdaki iki denklem kullanılarak, örneklem moment koşulları aşağıdaki gibidir.

$$\frac{1}{n}\sum(Y_i-\hat{b}_0-\hat{b}_1X_i)=0$$

$$\frac{1}{n}\sum Z_i(Y_i-\hat{b}_0-\hat{b}_1X_i)=0$$

Bu denklemlerin çözümü ile araç değişken (AD) tahmin edicileri elde edilir.

$$\hat{b}_1=\frac{n\sum Z_iY_i-\sum Z_i\sum Y_i}{n\sum Z_iX_i-\sum Z_i\sum X_i}=\frac{\sum(Z-\bar{Z})(Y_i-\bar{Y})}{\sum(Z-\bar{Z})(X_i-\bar{X})}=\frac{\sum z_iy_i}{\sum z_ix_i}$$

$$\hat{b}_0=\bar{Y}-\hat{b}_1\bar{X}$$

Bu yeni tahmin ediciler aşağıdaki özelliklere sahiptir:

- Z değişkeni dışsal yani $E(Z,u)=0$ ise, tahmin ediciler tutarlıdır.
- Büyük örneklemlerde araç değişken tahmin edicileri, yaklaşık normal dağılıma sahiptir. Basit regresyon modeli için

$$\hat{b}_1 \sim N\left(\beta_1,\frac{\sigma^2}{r_{XZ}^2\sum x_i^2}\right)$$

Burada r_{XZ}^2, araç değiken Z ve rassal bağımsız değişken X arasındaki kareli örneklem korelasyonudur.

- Hata teriminin varyans tahmin edici ise aşağıdaki gibidir.

$$\hat{\sigma}_{IV}^2=\frac{\sum(Y_i-\hat{b}_0-\hat{b}_1X_i)^2}{n-2}=\frac{\sum\hat{u}^2}{n-2}$$

2.4.Güçlü Araç Değişkeni Kullanmanın Önemi

Basit regresyon modelinde ($Y_i=\beta_0+\beta_1X_i+u_i$) $\hat{b}_1$ nin varyans tahmincisi

$$Var(\hat{b}_1)=\frac{\sigma^2}{r_{XZ}^2\sum x_i^2}$$

paydada, araç değişken Z ile içsel değişken X arasındaki kareli korelasyonu içermektedir. Eğer X ve u korelasyonsuz ise, en küçük kareler hala bir seçenektir. Böylece iki tahmin edicinin (EKK ve AD) etkinliğini karşılaştırabiliriz. β_1''nin araç değişkenler tahmin edicisinin varyansını aşağıdaki şekildeki gibi yazabiliriz.

$$Var(\hat{b}_1)=\frac{\sigma^2}{r_{XZ}^2\sum x_i^2}=\frac{Var(\hat{\beta}_1)}{r_{XZ}^2}$$

$r_{XZ}^2<1$ olduğundan dolayı araç değişkenler tahmin edicisinin varyansı her zaman en küçük kareler tahmin edicisinin varyansından daha büyük olacaktır ve böylece daha az *etkin* dir. Gerekli olmadığı sürece araç değişkenler tahmin prosedürünü kullanmak, en küçük kareler tahminine göre geniş güven aralıklarına ve daha az hassas çıkarıma yol açmaktadır. Z ve X arasındaki korelasyon 0.1 ise, araç değişkenler tahmin edicinin varyansı en küçük kareler tahmin edicinin varyansından 100 kez daha büyüktür. Eğer Z ve X arasındaki korelasyon 0.5 ise o zaman araç değişkenler tahmin edicinin varyansı en küçük kareler tahmin edicinin varyansından dört kez daha büyüktür.

Son yıllarda araç, içsel değişken X ile zayıf korelasyonlu olduğu zaman araç değişkenler tahmin edicinin davranışı üzerine pek çok araştırma var olmuştur. Zayıf bir araç kullanıldığı zaman, araç değişkenler tahmin edici büyük örneklemlerde bile çok eğilimli olabilir ve onun dağılımı yaklaşık olarak normal değildir. Böylece güven aralıkları ve hipotez testleri güvenilir sonuç vermez. Araç değişkenler zayıf olduğu durumda, tahmin güvenilir olmayacaktır.

2.5. Çok Değişkenli Regresyon Modelinde Araç Değişkenler Tahmini

Çok değişkenli regresyon modeli

$$Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_{k-1} X_{k-1} + u_i$$

Modelde *k* sayıda parametre, *k-1* sayıda bağımsız değişken vardır. Bağımsız değişkenlerden X_{k-1} 'in hata terimi ile korelasyonlu bir içsel değişken olduğunu biliyor veya şüpheleniyor olduğumuzu varsayalım. İlk $k-2$ değişken $(X_0 = 1, X_1, X_2, \ldots, X_{k-2})$ rassal hata terimi u ile korelasyonsuz olan dışsal değişkenlerdir. Araç değişkenler tahmini, her adımda en küçük kareler yönteminin kullanıldığı iki-aşamalı bir süreçtir.

Birinci aşamada sol tarafta içsel değişken X_{k-1} ve sağ tarafta bütün dışsal ve araç değişkenlerin yer aldığı regresyon modeli kurulur. L sayıda "dış" araç değişkenlere ($Z_1, Z_2, \ldots, Z_L$) sahip olduğumuz varsayımı altında, regresyon modeli aşağıdaki gibidir.

$$X_{k-1} = \gamma_0 + \gamma_1 X_1 + \gamma_2 X_2 + \cdots + \gamma_{k-2} X_{k-2} + \theta_1 Z_1 + \theta_2 Z_2 + \cdots + \theta_L Z_L + v$$

Burada v, sağ taraftaki bütün değişkenler ile korelasyonsuz olan bir rassal hata terimidir. Birinci aşamada nn küçük kareler yöntemi ile yukarıdaki model tahmin edilir ve X_{k-1}'in tahmni değeri bulunur.

$$\hat{X}_{k-1} = \hat{\gamma}_0 + \hat{\gamma}_1 X_1 + \hat{\gamma}_2 X_2 + \cdots + \hat{\gamma}_{k-2} X_{k-2} + \hat{\theta}_1 Z_1 + \hat{\theta}_2 Z_2 + \cdots + \hat{\theta}_L Z_L$$

X_{k-1}'in tahmni değeri ($\hat{X}_{k-1}$), modelde yer alan bütün dışsal ve araç değişkenlerin ağırlıklandırılmış bir ortalaması veya doğrusal bir kombinasyonudur.

İkinci aşamada regresyonda $\hat{X}_{k-1}$ ile X_{k-1}'yı yer değiştirilir ve model yeniden yazılır.

$$Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_{k-1}\hat{X}_{k-1} + u_i^*$$

burada u_i^*, rassal hata terimidir. Büyük örneklemlerde u_i^*, $\hat{X}_{k-1}$ dahil olmak üzere bağımsız değişkenler ile korelasyonsuz olduğundan dolayı, yukarıdaki model en küçük kareler yöntemi ile tahmin edilebilir. Bu denklemden en küçük kareler tahmin ediciler, $\hat{b}_0, \hat{b}_1, \hat{b}_2, \ldots, \hat{b}_{k-1}$ **araç değişkenler (*AD*)** tahmin edicileridir ve en küçük kareler yönteminin iki aşamada uygulaması ile bulunabildiklerinden dolayı, **iki-aşamalı en küçük kareler (*2AEKK*)** tahmin edicileri olarak adlandırılır. Bu tahminciler *AD* veya *2AEKK* veya *AD/2AEKK* tahmin ediciler olarak adlandırılır. Sağ tarafta bir içsel değişkenden daha fazlası olan genel durumda da aşamalar benzerdir.

$$Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_{k-1}\hat{X}_{k-1} + u_i^*$$

Yukarıdaki modelin en küçük kareler tahmin edicisi için, tahmin edici varyansları ve kovaryansları için standart formüller bir düzeltme ile kullanılabilir. Hata teriminin varyansının *AD/2AEKK* tahmin edicisi

$$\hat{\sigma}_{IV}^2 = \frac{\sum (Y_i - \hat{b}_0 - \hat{b}_1 \mathrm{X}_1 - \hat{b}_2 X_2 \ldots - \hat{b}_{k-1} X_{k-1})^2}{n-k}$$

orijinal model $Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_{k-1} X_{k-1} + u_i$ kalıntılarına dayanmaktadır.

Not: Ekonometrik yazılım programı kullanılıyorsa, iki-aşamalı en küçük kareler / araç değişkenler tahmini seçeneği seçildiği takdirde, ekonometrik yazılım otomatik olarak doğru varyans tahmin ediciyi kullanacaktır.

AD/2AEKK tahmin edilmiş $Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_{k-1}\hat{X}_{k-1} + u_i^*$'deki standart hataları kullanılarak büyük örneklemlerde geçerli olan parametrelerin t −testleri uygulanabilir ve aralık tahminlerini oluşturulabilir. Ayrıca, araç değişkenler zayıf değil ise, büyük örneklemlerde bileşik hipotez testleri geçerlidir.