Zaman Serisi Verileriyle Regresyon Analizi

Hüseyin Taştan[1]

[1]Yıldız Teknik Üniversitesi
Iktisat Bölümü

Textbook:
Introductory Econometrics (4th ed.)
J. Wooldridge

14 Mart 2013

Zaman Serisi Verileriyle Regresyon Analizi

Bu bölümde zaman serilerinin kullanıldığı doğrusal regresyon modellerinde OLS tahmin edicilerinin özellikleri incelenecektir.

- Kesitler-arası ve zaman serisi verileri arasındaki farklar
- Yaygın olarak kullanılan zaman serileri modellerinden örnekler
- Klasik varsayımlar altında OLS tahmin edicilerinin sonlu örneklem özellikleri
- Hipotez testleri ve çıkarsama
- Zaman serilerinde trent ve mevsimsellik (seasonality)

Zaman Serisi Verilerinin Özellikleri

- Zaman serilerinde veriler, kesitler-arası veriden farklı olarak, belli bir zaman sıralaması izlemektedir.
- Zaman serileri analizinde geçmişin geleceği etkilediği, ancak bunun tersinin geçerli olmadığı unutulmamalıdır.
- Kesitler-arası veride örneklem ilgili anakütleden rassal örnekleme ile elde ediliyordu. Farklı örneklemler farklı gerçekleşmeler üreteciğinden, bu örneklemler yardımıyla ulaşılan OLS tahmin değerleri de farklılık sergileyebiliyordu. Bu yüzden OLS tahmin edicilerini rassal değişken olarak değerlendirebiliyorduk.
- Bu durumda zaman serilerindeki rassallığı (randomness) nasıl yorumlayacağız?
- Zaman serisi değişkenlerinin (GSMH, İMKB indeksi, vs) bir sonraki dönemde hangi değerleri alacaklarını öngöremediğimiz için bu değişkenleri rassal değişken olarak düşünebiliriz.

ABD'de Enflasyon ve İşsizlik Oranlarına Dair Kısmi Zaman Serisi Verisi

Table 10.1

Partial Listing of Data on U.S. Inflation and Unemployment Rates, 1948–1996

Year	Inflation	Unemployment
1948	8.1	3.8
1949	−1.2	5.9
1950	1.3	5.3
1951	7.9	3.3
⋮	⋮	⋮
1994	2.6	6.1
1995	2.8	5.6
1996	3.0	5.4

Zaman Serisi Süreci ya da Stokastik Süreç

Tanım: Stokastik Süreç / Zaman Serisi Süreci

Zaman (t) endeksi taşıyan rassal değişkenlerin oluşturduğu diziye (sequence) **stokastik süreç** (stochastic process) ya da **zaman serisi süreci** (time series process) denir.

- Stokastik sözcüğü rassal (random) ile aynı anlamda kullanılmaktadır.
- Mevcut bir zaman serisi, stokastik sürecin olası bir gerçekleşmesi olarak görülebilir.
- Zamanda geriye gidip başka bir gerçekleşme elde edemeyeceğimiz için zaman serileri tek bir gerçekleşmenin sonuçlarıdır.
- Şüphesiz ilgilendiğimiz zaman serisi farklı tarihsel koşullar altında farklı olacaktı.
- Dolayısıyla, bir zaman serisi sürecinin bütün olası gerçekleşmelerinin oluşturacağı küme, burada, kesitler-arası verideki anakütlenin (population) rolünü üstlenecektir.

Zaman Serisi Modellerine Örnekler

Şimdi de sosyal bilimlerde kullanılan bazı modellere değinelim.

Statik Model

Eşanlı (contemporaneously) zaman endeksi taşıyan iki zaman serisi olsun : y ve z. y'yi z ile ilişkilendiren statik bir model şöyle yazılabilir:

$$y_t = \beta_0 + \beta_1 z_t + u_t \quad t = 1, 2, ..., n$$

Statik model, değişkenlerin birinci farkları (değişmeyi gösterir) arasında da formüle edilebilir:

$$\Delta y_t = \beta_1 \Delta z_t + u_t \quad t = 1, 2, ..., n$$

Statik Phillips Eğrisi

- Statik Phillips eğrisini statik modele bir örnek olarak verebiliriz:

Statik Phillips Eğrisi

$$inf_t = \beta_0 + \beta_1 unemp_t + u_t$$

- inf_t: enflasyon oranı ve $unemp_t$: işsizlik oranı
- Bu model **doğal işsizlik oranı** (natural rate of unemployment) ve **enflasyon beklentisinin** sabit olduğunu varsayar.
- inf_t ile $unemp_t$ değişkenleri arasındaki **eşanlı ödünümü** (contemporaneous tradeoff) bu model aracılığıyla inceleyebiliriz.

Sonlu Dağıtılmış Gecikme Modelleri (Finite Distributed Lag Models, FDL models)

- Sonlu dağıtılmış gecikme modellerinde (FDL models) y'yi, yani bağımlı değişkeni, belli bir gecikme (lag) ile etkileyen bir çok değişken mevcuttur:

Vergi Muafiyetinin Doğurganlığa Etkisi

$$gfr_t = \alpha_0 + \delta_0 pe_t + \delta_1 pe_{t-1} + \delta_2 pe_{t-2} + u_t$$

- gfr_t: doğurganlık oranı (fertility rate), doğurganlık yaşındaki 1000 kadına düşen bebek sayısı.
- pe_t: çocuk sahibi olmayı özendirmek için getirilen reel vergi muafiyeti.

Sonlu Dağıtılmış Gecikme Modelleri (FDL models)

- Vergi muafiyetinin doğurganlığa etkisini ele alan bu model ikinci sıradan bir FDL modeline örnektir.

İkinci sıradan bir FDL modeli

$$y_t = \alpha_0 + \delta_0 z_t + \delta_1 z_{t-1} + \delta_2 z_{t-2} + u_t$$

- Şimdi de z'de t döneminde bir kerelik (geçici) bir birimlik bir değişme olduğunu düşünelim:

$$..., z_{t-2} = c, \quad z_{t-1} = c, \quad z_t = c + 1, \quad z_{t+1} = c \quad z_{t+2} = c, ...$$

- Bu değişmenin y'de yaratacağı ceteris paribus etkiye, **etki çarpanı ya da etki çoğaltanı** denir. (impact multiplier or impact propensity).

Etki Çarpanının (Impact Multiplier) Hesaplanması

- İkinci sıradan bir FDL modelinde etki çarpanının hesaplanması.

$$y_{t-1} = \alpha_0 + \delta_0 c + \delta_1 c + \delta_2 c,$$

$$y_t = \alpha_0 + \delta_0 (c + 1) + \delta_1 c + \delta_2 c,$$
$$y_{t+1} = \alpha_0 + \delta_0 c + \delta_1 (c + 1) + \delta_2 c,$$
$$y_{t+2} = \alpha_0 + \delta_0 c + \delta_1 c + \delta_2 (c + 1),$$
$$y_{t+3} = \alpha_0 + \delta_0 c + \delta_1 c + \delta_2 c,$$

- İlk iki denklemden $y_t - y_{t-1} = \delta_0$ olduğu rahatlıkla görülebilir.
- Burada δ_0, t döneminde (cari) z'deki bir birim artışın y üzerindeki **ani** etkisini gösterir.
- δ_0, etki çarpanı olarak adlandırılır.

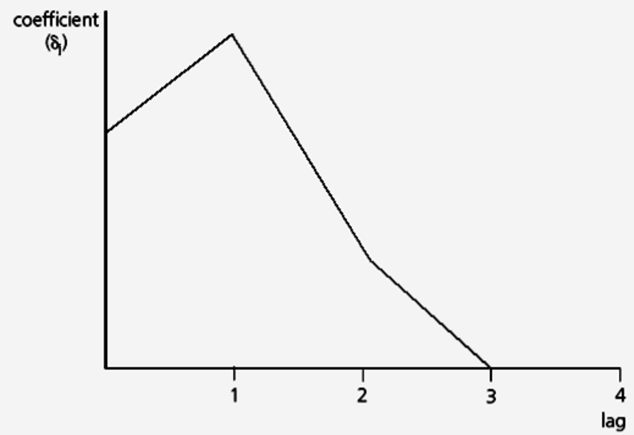
Etki Çarpanının (Impact Multiplier) Hesaplanması

- Benzer şekilde y'de geçici değişmenin olduğu t döneminden bir dönem sonraki değişme $y_{t+1} - y_{t-1} = \delta_1$'e, iki dönem sonraki değişme de $y_{t+2} - y_{t-1} = \delta_2$'ye eşit olacaktır.
- $t + 3$ döneminde y eski değerine geri dönecektir: $y_{t+3} = y_{t-1}$
- Bunun nedeni şu an incelenen modelin sadece iki dönem gecikme barındıran ikinci sıradan bir FDL modeli olmasıdır.
- δ_j'lerin j endeksine göre çizilen grafiği gecikme dağılımını (lag distribution) verecektir.
- Bu grafik, z'de meydana gelen geçici (temporary) bir artışın y üzerindeki dinamik etkisini gösterecektir.

Gecikme Dağılımı (Lag Distribution)

Figure 10.1

A lag distribution with two nonzero lags. The maximum effect is at the first lag.



Uzun Dönem Çoğaltanı (Long Run Propensity, LRP)

- z'deki sürekli (permanent) artışların y üzerindeki etkisini de bilmek isteriz.
- z, t döneminden önce $z = c$ (sabit), t döneminden itibaren ise $z = c + 1$ olsun.
- Artık t döneminden başlamak üzere z'de kalıcı bir değişiklik söz konusudur.

$$y_{t-1} = \alpha_0 + \delta_0 c + \delta_1 c + \delta_2 c,$$
$$y_t = \alpha_0 + \delta_0 (c+1) + \delta_1 c + \delta_2 c,$$
$$y_{t+1} = \alpha_0 + \delta_0 (c+1) + \delta_1 (c+1) + \delta_2 c,$$
$$y_{t+2} = \alpha_0 + \delta_0 (c+1) + \delta_1 (c+1) + \delta_2 (c+1),$$

Uzun Dönem Çoğaltanı (Long Run Propensity, LRP)

- z'deki kalıcı/sürekli bir artış, bir dönem sonra y'nin $\delta_0 + \delta_1$ kadar artmasına yol açar. İki dönem sonra ise y'deki artış $\delta_0 + \delta_1 + \delta_2$ kadardır. İki dönemden sonra ise y'de daha fazla artış meydana gelmez.
- Böylece cari ve gecikmeli z'lerin katsayılarının toplamı, z'deki kalıcı bir değişmenin y üzerindeki uzun dönemli etkisini gösterir.
- Bu uzun dönemli etkiye, **uzun dönem çarpanı ya da uzun dönem çoğaltanı (long run multiplier or long run propensity, LRP)** denir.
- Sonlu dağıtılmış gecikme (FDL) modellerinde, uzun dönem çoğaltanı ilginin ana odağıdır.

Uzun Dönem Çoğaltanı (Long Run Propensity, LRP)

Vergi Muafiyetinin Doğurganlığa Etkisi

$$gfr_t = \alpha_0 + \delta_0 pe_t + \delta_1 pe_{t-1} + \delta_2 pe_{t-2} + u_t$$

- Yukarıdaki modelde δ_0, pe'de 1 dolarlık artışın doğurganlık oranında yaratacağı eşanlı değişmeyi ölçer. Bu etki, davranışsal ve biyolojik nedenlerden ötürü, ya sıfır ya da çok küçük olacaktır.
- δ_1 ve δ_2, sırasıyla, bir dönem ve iki dönem önceki 1 dolarlık pe değişmelerinin etkilerini ölçmektedir. Bu katsayıların pozitif olmalarını bekleyebiliriz.
- Eğer pe, t döneminden itibaren kalıcı olarak $c + 1$ olursa, gfr, iki dönem sonra $\delta_0 + \delta_1 + \delta_2$ kadar artacaktır. İki dönem sonrası ise gfr'de değişme olmayacaktır.

q.cu Sıradan Sonlu Dağıtılmış Gecikme Modeli FDL(q)

- q.cu sıradan sonlu bir dağıtılmış gecikme modeli (a finite distributed lag model of order q):

FDL(q)

$$y_t = \alpha_0 + \delta_0 z_t + \delta_1 z_{t-1} + ... + \delta_q z_{t-q} + u_t$$

- FDL modelleri, bağımsız değişkenin (z) bağımlı değişken (y) üzerinde gecikmeli etkisinin (lagged effects) olup olmadığını görmemize yarar.
- Cari dönem değişkeni z_t'nin katsayısı, δ_0, etki çoğaltanı (impact multiplier or impact propensity) adını alır.
- Uzun dönem çoğaltanı (long-run propensity, LRP) tüm z_{t-j} katsayılarının toplamıdır:

$$LRP = \delta_0 + \delta_1 + ... + \delta_q$$

q.cu Sıradan Sonlu Dağıtılmış Gecikme Modeli

- z'nin gecikmeli değerleri arasında çoğu kez yüksek korelasyon bulunur. Bu durum çoklu-bağıntıya yol açar. Çoklu-bağıntı da δ_j'lerin ayrı ayrı kesin bir şekilde tahmin edilmelerini güçleştirir.
- FDL modellerinde birden fazla açıklayıcı değişken gecikmeli olarak bulunabilir.
- Cari dönem (contemporaneous) değişkenleri de, x_t ve w_t gibi, FDL modellerine eklenebilir.
- **Soru:** Aşağıda yıllık verilerle tahmin edilmiş FDL(2) modelinin etki ve uzun dönem çoğaltanlarını (LRP) bulun ve yorumlayın.

FDL(2)

$$\widehat{\mathrm{faiz}_t} = 1.6 + 0.48enf_t - 0.15enf_{t-1} + 0.32enf_{t-2}$$

- enf_t: enflasyon oranı ve $faiz_t$: faiz oranı

Klasik Varsayımlar Altında OLS Tahmin Edicilerinin Sonlu (Küçük) Örneklem Özellikleri

- Zaman serilerinde OLS tahmin edicilerinin sapmasız (unbiased) olabilmesi için üç varsayıma ihtiyaç vardır:

Varsayım $TS.1$: Parametrelerde Doğrusallık

$\{(x_{t1}, x_{t2}, ..., x_{tk}, y_t) : t = 1, 2, ..., n\}$ stokastik süreci aşağıdaki doğrusal modeli izler:

$$y_t = \beta_0 + \beta_1 x_{t1} + ... + \beta_k x_{tk} + u_t$$

Burada $\{u_t : t = 1, 2, ..., n\}$ hata terimi dizisidir. n gözlem sayısını (zaman periyodunu) gösterir.

- Bu varsayım kesitler-arası regresyondaki $MLR.1$'e denk gelmektedir.

Klasik Varsayımlar Altında OLS Tahmin Edicilerinin Sonlu (Küçük) Örneklem Özellikleri

- Zaman serilerinde x_{tj} değişkeninin iki endeksi vardır.
- t zaman endeksidir, j ise x'in no'sudur.
- FDL modellerinde her bir gecikmeli değişken ayrı bir x olarak tanımlanabilir:

$$x_{t1} = z_t, \quad x_{t2} = z_{t-1} \quad ve \quad x_{t3} = z_{t-2}$$

- Açıklayıcı değişkenlerin oluşturduğu kümeyi x_t ile göstereceğiz.
- x'lerden oluşan veri matrisi ise **X** olacaktır.
- **X**, $n \times k$ boyutludur.
- **X** matrisinin t.ci satırı t dönemine ait x değerlerinden oluşur.

Açıklayıcı Değişken Matrisine Bir Örnek

Table 10.2

Example of **X** for the Explanatory Variables in Equation (10.3)

t	*convrte*	*unem*	*yngmle*
1	.46	.074	.12
2	.42	.071	.12
3	.42	.063	.11
4	.47	.062	.09
5	.48	.060	.10
6	.50	.059	.11
7	.55	.058	.12
8	.56	.059	.13

Varsayım $TS.2$: Sıfır Koşullu Ortalama

Varsayım $TS.2$: Sıfır Koşullu Ortalama

Tüm dönemler itibarıyla açıklayıcı değişkenlere koşullu olarak, her t için hata teriminin (u_t) beklenen değeri sıfırdır.

$$E(u_t \mid \mathbf{X}) = 0, \quad t = 1, 2, ..., n.$$

- Bu kritik varsayım şunu söylüyor: t dönemine ait hata terimi, u_t, her bir x ile tüm dönemler itibariyle ilişkisizdir.
- Bu varsayım koşullu beklenen değer cinsinden ifade edildiği için, y ile x'lerin arasındaki ilişkinin biçiminin (form) doğru olarak belirlenmesi gerekmektedir.
- Yani, spesifikasyon hatası yapılmaması lazım.
- Eğer u_t, $\mathbf{X}$'den bağımsız ve $E(u_t) = 0$ ise, varsayım $TS.2$ otomatik olarak sağlanır.

Varsayım $TS.2$: Sıfır Koşullu Ortalama

- u_t'lerin aynı zamanda t dönemine ait x'lerle de ilişkisiz olması gerekmektedir:

$$E(u_t \mid x_{t1}, x_{t2}, ..., x_{tk}) = E(u_t \mid x_t) = 0$$

- Bu koşul u_t ve açıklayıcı değişkenlerin cari dönem itibariyle de ilişkisiz olduğunu ifade eder:

Cari Dönem Dışsallığı

$$Corr(x_{tj}, u_t) = 0$$

Her j için geçerlidir.

- $TS.2$ varsayımı cari dönem dışsallığından (contemporaneous exogeneity) öte koşullar getirmektedir.
- u_t, x_{sj} ile, $s \neq t$ iken bile ilişkisiz olmalıdır.
- Yani, t dönemi hata terimi u_t'nin ortalama değeri, x'lerin tüm geçmiş, şimdiki (cari) ve gelecek değerleriyle ilişkisiz olmalıdır.

Varsayım $TS.2$: Sıfır Koşullu Ortalama

- $TS.2$ sağlandığında x'lerin **kesin olarak dışsal** (strictly exogenous) olduğunu söyleriz.
- OLS'nin tutarlılığı için cari dışsallığın sağlanması yeterlidir. Ancak, OLS'nin sapmasızlığı için kesin dışsallık gerekmektedir (bkz. Bölüm 11).
- Kesitler-arası regresyonda örneklemin rassal oluşu $(MLR.2)$ kesin dışsallık varsayımını gereksiz kılıyordu.
- Zaman serilerinde rassal örnekleme olmadığı için kesin dışsallık varsayımına ihtiyaç duyuyoruz.

Varsayım $TS.2$: Sıfır Koşullu Ortalama

- Varsayım $TS.2$, x'lerin ve u'nun kendi geçmişleri ile korelasyonlarına izin vermektedir.
- İzin verilmeyen, u_t'nin beklenen değerinin x'lerle zaman içinde ileri ve geriye doğru ilişkili olmasıdır.
- $TS.2$'nin sağlanamamasına yol açan başlıca iki faktör **ihmal edilmiş değişkenler** ve **ölçme hataları**dır.
- Ancak başka nedenler de varsayımın ihlaline yol açabilmektedir.

Varsayım $TS.2$: Sıfır Koşullu Ortalama

- Şu basit statik, yani açıklayıcı değişkenler arasında gecikmeli değişkenin olmadığı regresyonu ele alalım :

$$y_t = \beta_0 + \beta_1 z_t + u_t$$

- $TS.2$ varsayımı, sadece u_t ve z_t'nin ilişkisiz olmasını gerektirmiyor.
- u_t'nin, z_t'nin tüm geçmiş $\{z_{t-1}, z_{t-2}, ...\}$ ve gelecek $\{z_{t+1}, z_{t+2}, ...\}$ değerleri ile de ilişkisiz olmasını koşul olarak koyuyor.
- Bu koşulun iki sonucu (implication) vardır :
  1. z'nin y üzerinde gecikmeli etkisi **(lagged effect)** yoktur.
  2. Eğer gecikmeli etkisi varsa, FDL modeli tahmin etmemiz gerekmektedir.
- **Kesin dışsallık** (strict exogeneity) varsayımı, u'da t anında oluşacak bir değişmenin z'nin gelecek değerlerine etki etmeyeceğini varsayar.
- Bu ise, y'den z'nin gelecek değerlerine bir etkinin (feedback) olmadığı anlamına gelir.

Varsayım $TS.2$: Sıfır Koşullu Ortalama

- Şehirlerde işlenen cinayet sayılarını (nüfusa oran olarak), nüfus başına düşen polis sayısı ile açıklayalım :

$$mrdrte_t = \beta_0 + \beta_1 polpc_t + u_t$$

- u_t'nin $polpc_t$ ile ilişkisiz olmasını varsaymamız makuldur.
- Hatta $polpc_t$'nin geçmiş değerleri ile de ilişkisiz olduğunu varsaymamız da çok sakıncalı olmaz ve böyle olduğunu varsayalım.
- Diyelim ki, şehir yönetimi polis sayılarını geçmiş cinayet sayılarına bakarak değiştiriyor.
- Bu durumda, $mrdrte_t \rightarrow polpc_{t+1}$ yönünde bir etkileşim olacaktır.
- Bu ise, $u_t \rightarrow polpc_{t+1}$ etkileşimi olduğu anlamına gelecektir. Bu durumda $TS.2$ varsayımı ihlal edilmiş olacaktır.

Varsayım $TS.2$: Sıfır Koşullu Ortalama

- Distributed lag modellerinde (DLM) u_t'nin z'nin geçmiş değerleriyle ilişkili olması sorun olmaz, zira, modelde z'nin geçmiş değerlerini, $\{z_{t-1}, z_{t-2}, ...\}$, açıklayıcı değişken olarak zaten kullanıyoruz (kontrol ediyoruz).
- Ancak, u'dan z'nin gelecek değerlerine doğru bir etki (feedback), $u_t \rightarrow z_{t+1}, z_{t+2}, ...$, her zaman sorun yaratacaktır.
- Kesin dışsal (strict exogeneity) olan açıklayıcı değişkenler y'nin geçmiş değerlerinden etkilenmez.
- Örneğin, t yılındaki yağmur miktarı, Y_t, bu yılın ve önceki yılların buğday üretiminden, $\{Q_t, Q_{t-1}, Q_{t-2}, ...\}$, etkilenmez.
- Bu aynı zamanda şu anlama da gelir : gelecek yılların yağmur miktarı, $\{Y_{t+1}, Y_{t+2}, ...\}$, bu yılın ve geçen yılların buğday üretiminden, $\{Q_t, Q_{t-1}, Q_{t-2}, ...\}$, etkilenmez.

Varsayım $TS.2$: Sıfır Koşullu Ortalama

- Ancak, tüm tarım girdileri yağmur gibi değildir. Örneğin, işgücü girdisini çiftlik sahibi geçen yılın hasılasına bakarak belirleyebilir.
- Dolayısıyla, işgücü kesin dışsal değildir.
- Sosyal bilimlerde kullandığımız pek çok açıklayıcı değişken böyledir.
- Para arzı artış hızı, sosyal refah harcamaları, yollardaki hız limitleri vs.
- Tüm bu değişkenler, çoğu zaman, kontrol değişkeni y'nin geçmişte aldığı değerlere bakılarak belirlenmektedirler, dolayısıyla da kesin dışsallık varsayımının ihlali söz konusudur.

Varsayım $TS.2$: Sıfır Koşullu Ortalama

- $TS.2$ varsayımı gerçekçi olmamasına rağmen OLS tahmin edicilerin sapmasız (unbiased) olmasını sağlamak için kullanılmaktadır.
- Çoğu zaman $TS.2$ varsayımı ondan daha katı olan şu varsayım ile değiştirilir:
- **Açıklayıcı değişkenler (x'ler) rassal (random) değildir ya da tekrarlanan örneklemlerde sabit (fixed) değerlerden ibarettirler**.
- Bu varsayım, otomatik olarak $TS.2$ varsayımını sağlamaktadır.
- Ancak, rassal-olmama (nonrandomness) varsayımının zaman serileri gözlemleri için doğru olamayacağı açıktır.
- Oysa $TS.2$ varsayımı x'lerin rassallık niteliğine dayandığı için daha gerçekçidir.
- Ancak, sapmasızlığın sağlanması için $x \leftrightarrow u$ ilişkilerinin nasıl olması gerektiği konusunda katı koşullar getirmektedir.

Varsayım $TS.3$: Tam Çoklu Bağıntının Olmaması

Varsayım $TS.3$: Tam Çoklu Bağıntının Olmaması

Örneklemde (dolayısıyla altında yatan zaman serisi sürecinde) bağımsız değişkenler sabit değildir ve diğer değişkenlerin tam bir doğrusal kombinasyonu olmamalıdır.

- OLS'nin sapmasızlığı için gerekli son varsayım standart **tam çoklu-bağıntı olmaması** varsayımıdır .
- x'ler arasında ilişki olabilir, ancak tam (perfect) korelasyona izin verilmemektedir.
- Ayrıca, x'lerde değişme olması, yani sabit olmamaları da bu varsayım içinde yer almaktadır.

OLS'nin Sapmasızlığı

Teorem 10.1: OLS'nin Sapmasızlığı

$TS.1 - TS.3$ varsayımları altında OLS tahmin edicileri X'e koşullu olarak sapmasızdır:

$$\mathsf{E}(\hat{\beta}_j) = \beta_j, \quad j = 0, 1, 2, \ldots, k$$

- Bu teoremin ispatı kesitler-arası regresyon için verilen Teorem 3.1'in ispatının aynıdır.
- Ancak, oradaki rassal örnekleme (random sampling) varsayımının yerini burada **'açıklayıcı değişkenlerin tüm zamanlar için değerleri kontrol edilmişken, her bir t dönemi için u_t sıfır ortalamaya sahiptir.'** varsayımı almıştır.
- Bu varsayım sağlanamazsa OLS'nin sapmasız olduğu kanıtlanamaz.
- İki varsayım daha ekleyerek zaman serileri için Gauss-Markov varsayımlarını tamamlayalım : **Sabit varyans** (Homoscedasticity) ve **Ardışık bağımlılık** (serial correlation)

Varsayım $TS4$: Sabit varyans Varsayım $TS5$: Ardışık Bağımlılığın Olmaması

- İki varsayım daha ekleyerek zaman serileri için Gauss-Markov varsayımlarını tamamlayalım : **Sabit varyans** (Homoscedasticity) ve **Ardışık bağımlılık** (serial correlation) olmaması varsayımları

Varsayım $TS.4$: Sabit Varyans

X'e koşullu u_t'nin varyansı her t dönemi için sabittir.
$Var(u_t|\mathbf{X}) = \sigma^2, \quad t = 1, 2, , ..., n.$

Varsayım $TS.5$: Ardışık Bağımlılılığın olmaması

Her $t \neq s$ için, **X**'e koşullu biçimde farklı iki zaman dönemine ait hata terimleri arasındaki korelasyon sıfırdır: $Corr(u_t, u_s|\mathbf{X}) = 0$.

$TS.5$: Ardışık Bağımlılığın Olmaması

Varsayım $TS.5$: Ardışık Bağımlılığın olmaması

Her $t \neq s$ için, **X**'e koşullu biçimde farklı iki zaman dönemine ait hata terimleri arasındaki korelasyon sıfırdır: $Corr(u_t, u_s|\mathbf{X}) = 0$.

- $TS.5$ varsayımı iki ayrı zaman dönemine (t ve s diyelim) ait u'ların x'lere koşullu olarak ilişkisiz olması anlamına gelmektedir.
- x'ler rassal-olmayan (nonrandom) değişken ya da tekrarlanan örneklemlerde sabit değerler olarak ele alınırsa, bu halde x'lere koşullu olma kaydını kaldırırız:

$$Corr(u_t, u_s) = 0$$

- Bu durum sağlanmadığında hata terimi (u) **ardışık bağımlılık** (serial correlation) ya da **otokorelasyon** (autocorrelation) içeriyor demektir.

$TS.5$: Ardışık Bağımlılığın Olmaması

- Yani, hata terimleri zaman dönemleri itibariyle (across time) ilişkilidirler.
- Otokorelasyon, ard arda gelen u'ların tümünün birden pozitif ya da tümünün birden negatif olması şeklinde ortaya çıkar.
- Oysa, ideal durum bu hata teriminin tamamen birbirinden bağımsız olarak rasgele dağılmaları durumu idi.
- Kesitler-arası regresyonda bu varsayımı yapmadık. Nedeni ise rassal örnekleme varsayımıydı.
- Rassal örnekleme varsayımı altında herhangi iki i ve h gözlemlerine ait hata terimleri, u_i ve u_h, birbirinden bağımsızdır. Bu, tüm açıklayıcı değişkenlere koşullu olarak da böyledir.
- Demek ki otokorelasyon sadece zaman serileri regresyonlarına özgü bir sorundur.
- Ancak, örneklemin rassal olmadığı kesitler-arası verilerde de otokorelasyon sorunu çıkabilir. Hata terimlerinin şehirler itibariyle ilişkili olması gibi.

OLS Örnekleme Varyansları

Teorem 10.2: OLS Örnekleme Varyansları

Zaman serileri Gauss-Markov varsayımları ($TS.1 - TS.5$) altında

$$\text{Var}(\hat{\beta}_j \mid \mathbf{X}) = \frac{\sigma^2}{SST_j(1-R_j^2)}, \quad j = 1, 2, \ldots, k$$

Burada

$$SST_j = \sum_{t=1}^{n} (x_{tj} - \bar{x}_j)^2$$

x_j'deki örneklem değişkenliği, R_j^2 ise x_j'nin diğer tüm x değişkenlerine (sabit terim içeren) regresyonundan elde edilen belirlilik katsayısıdır.

Bu, Gauss-Markov koşulları altında kesitler-arası regresyon için türettiğimiz varyans ile aynı şeydir. Çoklu-bağıntı gibi varyansın büyük çıkmasına sebep olan faktörler aynı etkiyi göstermektedir.

Gauss Markov Teoerimi

Teorem 10.3: $\hat{\sigma}^2$'nin Sapmasız Tahmini

$TS.1 - TS5$ varsayımları altında tahmin edici $\hat{\sigma}^2 = \frac{SSR}{dof}$, σ^2'nin sapmasız tahmin edicisidir. Burada serbestlik derecesi: $dof = n - (k+1)$

Hata terimleri varyansının OLS tahmin edicisi, $TS.1 - TS.5$ varsayımları altında sapmasızdır ve yine bu varsayımlar altında Gauss-Markov teoremi sağlanmaktadır.

Teorem 10.4: Gauss-Markov Teoremi

$TS.1 - TS5$ varsayımları altında OLS tahmin edicileri X'e koşullu en iyi doğrusal sapmasız tahmin eidicilerdir (BLUE, DESTE).

Klasik Doğrusal Model Varsayımları Altında Çıkarım

- OLS tahmin edicileri, $TS.1 - TS.5$ varsayımları altında tıpkı $MLR.1 - MLR.5$ varsayımları altında olduğu gibi arzu edilir küçük (sonlu) örneklem özelliklerine sahip olmaktadırlar.
- Zaman serileri regresyonlarında hipotez testleri yapabilmek ve güven aralıkları oluşturabilmek için, başka bir deyişle, standart hata, t ve F testlerini kullanabilmemiz için kesitler-arasıdaki varsayımının bir benzerini burada da yapacağız:

Varsayım $TS.6$: Normallik

u_t hata terimleri X'ten bağımsızdır ve bağımsız ve özdeş(normal) dağılımlıdır (iid). Yani $u \sim N(0, \sigma^2)$.

- $TS.6$ varsayımı, $TS.3$, $TS.4$ ve $TS.5$ varsayımlarının geçerli olmalarını zorunlu kılar.
- Başka bir deyişle, $TS.6$ sağlanmışsa, $TS.3$, $TS.4$ ve $TS.5$ otomatik olarak sağlanmış olur.
- Ancak, bağımsızlık ve normallik varsayımları dolayısıyla $TS.6$ daha katı bir varsayımdır.

Klasik Doğrusal Model Varsayımları Altında Çıkarım

Teorem 10.5: Normal Örnekleme Dağılımları

$TS.1 - TS.6$ varsayımları altında, zaman serileri için klasik varsayımlar, OLS tahmin edicilerinin X'e koşullu dağılımı normaldir. Sıfır hipotezi altında t istatistikleri t dağılımına, F istatistikleri F dağılımına uyar. Güven aralıkları standart biçimde oluşturulabilir.

Klasik Doğrusal Model Varsayımları Altında Çıkarım

- Teorem 10.5, $TS.1 - TS.6$ varsayımları sağlandığında kesitler-arası regresyonda tahmin (estimation) ve çıkarımla (inference) ilgili olarak elde edilen tüm sonuçların zaman serileri regresyonuna da uygulanabileceğini ifade ediyor.
- Zaman serileri için klasik doğrusal model varsayımları $TS.1 - TS.6$, kesitler-arası regresyonu varsayımlarına kıyasla daha katı koşullar getirmektedir.
- Özellikle kesin dışsallık ve otokorelasyon olmaması varsayımları çoğu kez gerçekçi olmaktan uzak olabilirler.

Örnek 10.1: Statik Phillips Eğrisi

- Ortalamada işsizilik ve enflasyon arasında bir ödünümün varlığını araştırmak için
- $H_0 : \beta_1 = 0$ ile $H_1 : \beta_1 < 0$ test ediebilir.

$$inf_t = \beta_0 + \beta_1 unemp_t + u_t$$

- Klasik varsayımlar geçerliyse OLS t istatistiği kullanılabilir.

Statik Phillips Eğrisi (1948-1996) PHILLIPS.RAW

$$\widehat{\mathrm{inf_t}} = \underset{(1.72)}{1.42} + \underset{(0.289)}{0.468}\, unem_t$$

$$n = 49 \quad R^2 = 0.053 \; \bar{R}^2 = 0.033$$

Örnek 10.1: Statik Phillips Eğrisi

- $\hat{\beta}_1$ beklenilenin aksine $(+)$ işaretli çıkmıştır. Enflasyonla işsizlik arasında beklediğimiz zıt yönlü bir ödünüm (tradeoff) gözükmemektedir.
- Bu, kısmen modelin yetersizliğinden ileri gelebilir. Beklentilerin (expectations) ilave edildiği genişletilmiş (augmented) Phillips eğrisi modeli daha başarılı sonuç vermektedir.
- İkincisi, modelin tatmin edici olmaması klasik varsayımların sağlanamamasından da ileri gelebilir.

Örnek 10.2: Enflasyon ve Bütçe Açığının Faiz Oranı Üzerinde Etkileri (ABD)

- $i3$: üç aylık hazine bonosu faiz haddi
- inf : yıllık tüketici enflasyonu
- def : GSYİH'ya oran olarak federal bütçe açıkları.
- Dönem : 1948-1996.

Enflasyon ve Bütçe Açığının Faiz Oranı Üzerinde Etkileri (INTDEF.RAW)

$$\widehat{\mathrm{i3}_t} = \underset{(0.44)}{1.25} + \underset{(0.076)}{0.613}\, inf_t + \underset{(0.118)}{0.700}\, def_t$$

$$n = 49 \quad R^2 = 0.697 \; \bar{R}^2 = 0.683$$

Örnek 10.2: Enflasyon ve Bütçe Açığının Faiz Oranı Üzerinde Etkileri (ABD)

- Diğer her şey sabitken, enflasyonda yüzde 1 puanlık (one percentage point) artış kısa vadeli faizlerde 0.613 puanlık artış yaratıyor.
- Katsayıların t değerleri oldukça yüksektir, dolayısıyla istatistiksel olarak anlamlıdırlar.
- Tabii, klasik varsayımlar burada sağlanmışsa bu değerleri çıkarım (inference) amacıyla kullanabileceğiz.

Fonksiyonel Biçim ve Kukla Değişkenler

- Kesitler-arası regresyonda gördüğümüz tüm fonksiyonel biçimleri zaman serilerinde de kullanabileceğiz.
- Özellikle doğal logaritma zaman serilerinde çok kullanılmaktadır.
- Böylece değişkenlerin y üzerindeki etkisi, ölçü birimlerinden bağımsız olarak, **sabit yüzde** cinsinden elde edilebilmektedir.

Örnek 10.3: İstihdam ve Asgari Ücret (Porto Riko)

- $prepop$: istihdam oranı
- $mincov$ = [(ortalama asgari ücret)/ortalama ücret] × [(asgari ücretli sayısı)/(tüm çalışanlar)]
- Böylece, $mincov$ değişkeni asgari ücretin ortalama ücrete göre görece önemini ölçüyor.
- $usgdp$: ABD'nin GSMH'ı.
- Dönem : 1950-1987.

İstihdam ve Asgari Ücret (PRMINWGE.RAW)

$$\log(\widehat{\text{prepop}})_t = \underset{(0.77)}{-1.05} - \underset{(0.065)}{0.154}\, log(mincov)_t - \underset{(0.118)}{0.012}\, usgdp_t$$

$$n = 38 \quad R^2 = 0.661 \ \bar{R}^2 = 0.641$$

Örnek 10.3: İstihdam ve Asgari Ücret (Porto Riko)

- $prepop$'un $mincov$'a göre esnekliği -0.154 bulunmuştur ve anlamlıdır.
- Daha yüksek bir asgari ücret, istihdam oranını düşürmektedir. Bu, iktisat teorisinin öngörüsüne uygun bir sonuçtur.
- $usngp$ serisi anlamsız çıkmıştır. Ancak, ileride göreceğiz ki regresyona trent ekleyince bu sonuç değişecektir.

FDL Modeli ve Logaritmik Fonksiyonel Biçim

- Logaritmik fonksiyonel biçimi sonlu "distributed lag" modelleri (FDLM) için de kullanabiliriz.
- Örneğin, para talebini (M) GDP'nin bir fonksiyonu olarak çeyrek yıllık veriden şöyle tahmin edebiliriz:

FDL(4)

$$log(M_t) = \alpha_0 + \delta_0 log(GDP)_t + \delta_1 log(GDP)_{t-1} + \delta_2 log(GDP)_{t-2}$$

$$+\delta_3 log(GDP)_{t-3} + \delta_4 log(GDP)_{t-4} + u_t$$

FDL Modeli ve Logaritmik Fonksiyonel Biçim

- δ'ların toplamına, $\delta_0 + \delta_1 + \delta_2 + \delta_3 + \delta_4$, **uzun-dönem esnekliği** (long-run elasticity) denir.
- GDP'de yüzde 1'lik sürekli (permanent) artışın, para talebinde 4 çeyrek yıl sonra doğuracağı yüzde artışı gösterir.
- Burada cari dönem (t) değişkeninin katsayısı δ_0, 'the impact propensity' ya da **kısa dönem esnekliği** (short-run elasticity) adını alır.
- GDP'de yüzde 1'lik bir artışın para talebinde doğuracağı ani etkiyi yüzde olarak gösterir.

Zaman Serisi Analizi ve Kukla Değişkenler

- Kukla değişkenler zaman serilerinde çok sıkça kullanılır ve çok yararlıdırlar.
- Örneğin, herhangi bir A partisinin iktidarda olduğu yıllarda 1, olmadığı yıllarda 0 değerini alan bir kukla değişkeni tanımlayıp regresyonumuzda açıklayıcı değişken olarak kullanabiliriz.
- Savaş yılları, krizler, depremler vb. bazı özel dönemleri kukla değişken kullanarak izole edebiliriz.
- Örneğin, II.Dünya Savaşı yıllarında 1, diğer yıllarda 0 değerini alan bir kukla değişkeni yaratarak savaş yıllarının bağımlı değişken (y) üzerindeki özel etkisini ölçebiliriz.
- Belli olayların kukla değişken kullanarak etkilerini ölçmeye **event study** denir.

Örnek 10.4: Vergi Muafiyetinin Doğurganlık Üzerindeki Etkisi (ABD, 1913-1984)

- gfr: Doğurganlık dönemindeki 1000 kadın başına düşen doğum sayısı
- pe: Kişisel vergi muafiyeti (çocuk sahibi olunduğunda)
- $ww2$: ABD'nin II. Dünya Savaşına olduğu yıllar için 1 değeri alan kukla değişken
- $pill$: Doğum kontrol hapının kullanılmaya başlandığı 1963'den sonra 1 değeri alan kukla değişken

Vergi Muafiyeti ve Doğurganlık (FERTIL3.RAW)

$$\widehat{\text{gfr}_t} = \underset{(3.21)}{98.68} + \underset{(0.030)}{0.083}\, pe_t - \underset{(7.46)}{24.24}\, ww2_t - \underset{(4.08)}{31.59}\, pill_t$$

$$n = 72 \quad R^2 = 0.473 \;\; \bar{R}^2 = 0.450$$

Örnek 10.4: Vergi Muafiyetinin Doğurganlık Üzerindeki Etkisi (ABD, 1913-1984)

- Tüm katsayılar çift taraflı bir alternatif hipoteze karşı yüzde 1 düzeyinde anlamlıdırlar.
- II. Dünya savaşı yıllarında, ceteris paribus, doğurganlık çağındaki 1000 kadın başına düşen doğum sayısı 24 azalmıştır.
- gfr değişkeninin 65-127 arasında değiştiğini düşünürsek bu sayı oldukça önemlidir.
- Yine her şey aynı iken, doğum kontrol haplarının piyasaya çıktığı yıldan sonraki doğum sayılarında ciddi azalma olmuştur.
- Doğumları teşvik için getirilen vergi muafiyeti değişkeni pe'nin örnek ortalaması 100.4 dolar olup, minimum ve maksimum değeri sırasıyla 0 ve 243.8'dir.

Örnek 10.4: Vergi Muafiyetinin Doğurganlık Üzerindeki Etkisi (ABD, 1913-1984)

- Bu muafiyette yapılan 12 dolarlık bir artış 1000 kadın başına düşen doğum sayısını 1 adet artıracaktır (12 x 0.083 = 1).
- Doğurganlık, pe'deki değişmelere belli bir gecikme ile tepki (response) verebilir. Dolayısıyla, 2 yıl gecikmenin (lag) kullanıldığı bir DLM tahmin edelim:

Vergi Muafiyeti ve Doğurganlık: FDL(2)

$$\begin{aligned}\widehat{\text{gfr}_t} &= \underset{(3.28)}{95.87} + \underset{(0.126)}{0.073}\, pe_t - \underset{(0.1557)}{0.0058}\, pe_{t-1} \\ &+ \underset{(0.126)}{0.034}\, pe_{t-2} - \underset{(10.73)}{22.13}\, ww2_t - \underset{(3.98)}{31.30}\, pill_t\end{aligned}$$

$$n = 70 \quad R^2 = 0.499 \;\; \bar{R}^2 = 0.459$$

- Örneklem hacmi gecikme sayısı (2) kadar azaldı.

Örnek 10.4: Vergi Muafiyetinin Doğurganlık Üzerindeki Etkisi (ABD, 1913-1984)

- pe_t, pe_{t-1} ve pe_{t-2} arasında çok yüksek çoklu-bağıntı olduğu için katsayılarının standart hataları çok yüksek çıktı .
- Ayrı ayrı hiçbiri anlamlı değildir. Ancak, üçü birden (jointly) anlamlıdır (F istatiğinin p-değeri: 0.012).
- pe katsayılarının üçü de anlamsız çıktığı için pe'nin gfr üzerindeki etkisinin cari dönem itibariyle mi (contemporaneous effect) yoksa gecikmeli mi olduğunu bilemiyoruz.
- Bunun için pe_{t-1} ve pe_{t-2}'nin katsayılarının birlikte anlamlı olup olmadığını yine F testi ile test ediyoruz.
- Testin p-değeri 0.95 çıktığı için gecikmeli değişkenlerin etkisi yoktur diyoruz ve statik modeli tercih ediyoruz.

Örnek 10.4: Vergi Muafiyetinin Doğurganlık Üzerindeki Etkisi (ABD, 1913-1984)

- Bu FDL(2) modelinde uzun-dönem çoğaltanı (long-run propensity): $0.073 - 0.0058 + 0.034 = 0.101$ olarak buluyoruz.
- Bu tahminin anlamlılığını sınamak için standart hatasını bilmemiz gerek.
- Bunun için Bölüm 4.4'deki yola başvuruyoruz:
- $\theta_0 = \delta_0 + \delta_1 + \delta_2$ **uzun-dönem çoğaltanı**
- $\delta_0 = \theta_0 - \delta_1 - \delta_2$'ı modelde yerine koyarsak:

$$\begin{aligned} gfr_t &= \alpha_0 + \delta_0 pe_t + \delta_1 pe_{t-1} + \delta_2 pe_{t-2} + \dots \\ &= \alpha_0 + (\theta_0 - \delta_1 - \delta_2) pe_t + \delta_1 pe_{t-1} + \delta_2 pe_{t-2} + \dots \\ &= \alpha_0 + \theta_0 pe_t + \delta_1 (pe_{t-1} - pe_t) + \delta_2 (pe_{t-2} - pe_t) + \dots \end{aligned}$$

Örnek 10.4: Vergi Muafiyetinin Doğurganlık Üzerindeki Etkisi (ABD, 1913-1984)

- Son modeli tahmin ettiğimizde stderror$(\hat{\theta}_0) = 0.030$ ve böylece t istatistiğini 3.37 buluruz ve anlamlıdır.
- Her üç δ da tek tek anlamsız çıktığı halde onların toplamı olan uzun-dönem çoğaltanı kuvvetli bir şekilde sıfıra karşı testi geçmektedir.
- θ_0 için % 95'lik bir güven aralığı oluşturursak, bu aralığın $0.041 - 0.16$ olduğunu görürüz.

Zaman Serisi Analizinde Trent Kullanılması

- Pek çok ekonomik değişken zaman içinde artma eğilimi gösterir. Yani zaman içinde trent gösterir.
- İki değişkenin aynı ya da zıt yönde trent göstermesi onların mutlaka birbirleri üzerinde etkiye sahip oldukları anlamına gelmez.
- Herhangi iki seri çoğu kez diğer başka gözlenemeyen faktörlerin etkisiyle zaman içinde trent gösterdikleri için ilişkili çıkmaktadır.
- Değişkenlerdeki trendi hangi istatistiksel modellerle ifade edebiliriz?

Doğrusal Zaman Trendi Modeli

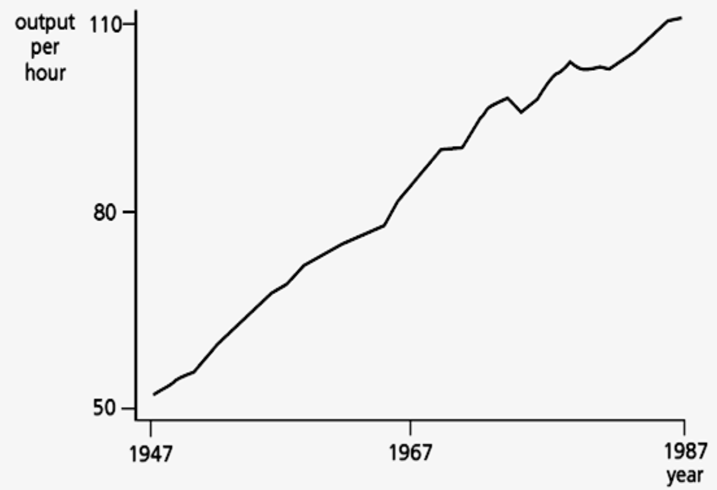
$$y_t = \alpha_0 + \alpha_1 t + e_t, \qquad t = 1, 2, \dots$$

- En yaygın model **doğrusal zaman trendi** (linear time trend) modelidir.

1947-87 ABD Saat Başına Çıktı: 1977=100

Figure 10.2

Output per labor hour in the United States during the years 1947–1987; 1977 = 100.



Zaman Serisi Analizinde Trent Kullanılması

- Burada, (e_t), **bağımsız, özdeş dağılmış** (independent, identically distributed, iid) bir silsiledir (sequence).

$$E(e_t) = 0 \qquad Var(e_t) = \sigma_e^2$$

- α_1, diğer faktörler (e_t'de içerilen) sabit iken, zaman endeksinde 1 dönemlik (period) bir değişme olduğunda y'de meydana gelecek değişmedir.

Zaman Serisi Analizinde Trent Kullanılması

- Burada, $\Delta e_t = 0$ iken,

$$\Delta y_t = y_t - y_{t-1}$$

- Diğer bir zaman trendi gösteren silsile, ortalaması zamanın bir fonksiyonu olan süreçlerdir:

$$E(y_t) = \alpha_0 + \alpha_1 t.$$

- $\alpha_1 > 0$ iken y_t zamanla artacak, $\alpha_1 < 0$ ise azalacaktır.
- Ancak burada doğrusal bir şekilde (bir çizgi üzerinde) artan ya da azalan y_t değil onun ortalamasıdır.
- Rassal değerler alan e_t, dolayısıyla y_t zig zaglar izleyecektir.

Zaman Serisi Analizinde Trent Kullanılması

- Bu durumda ortalamanın aksine varyans zaman içinde sabittir:

$$Var(y_t) = Var(y_{t-1}) = \sigma_e^2$$

- Eğer $\{e_t\}$ iid silsilesi ise $\{y_t\}$ bağımsız bir silsile olacak, ancak özdeş olmayacaktır.
- e_t de zaman içinde kendi geçmişi ile ilişkili olabilir. Bu durum daha gerçekçidir.

Zaman Serisi Analizinde Trent Kullanılması

- Çoğu iktisadi seri zaman içinde sabit bir ortalama hızla artar. Bu durumda trendi üstel olarak modellemek uygun olur.
- Eğer seri hep pozitif değerler alıyorsa, üstel trendi doğal logaritma kullanarak ifade edebiliriz:

Üstel Trent Modeli

$$y_t = exp(\beta_0 + \beta_1 t + e_t)$$

- β_1 büyüme oranıdır. İki tarafın logaritmasını alırsak:

$$log(y_t) = \beta_0 + \beta_1 t + e_t$$

$$\Delta log(y_t) \approx \frac{y_t - y_{t-1}}{y_{t-1}}$$

- $\Delta e_t = 0$ kabul edildiğinde

$$\Delta log(y_t) = \beta_1$$

Zaman Serisi Analizinde Trent Kullanılması

- **Karesel trent (quadratic trend):** Bir zaman serinin eğimi zaman içinde artıyorsa, yani büyüme giderek hızlanıyorsa (hiper enflasyonda olduğu gibi), kareli t terimi de regresyona eklenir:

Karesel Trent Modeli

$$y_t = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + e_t$$

- Eğer α_1 ve α_2 pozitifse, eğim t ile birlikte artacaktır.
- Eğer $\alpha_1 > 0$ ve $\alpha_2 < 0$ ise trent kambur şeklinde olacaktır ve eğim t ile birlikte azalacaktır.

$$\frac{\Delta y_t}{\Delta t} \approx \alpha_1 + 2\alpha_2 t$$

Trende Sahip Değişkenlerin Regresyonda Kullanımı

- Değişkenlerin trent içermesi $TS.1 - TS.6$ varsayımlarımızı bozmaz.
- Ancak, eğer y'ye trent kazandıran gözlenemeyen faktörler x'lerle de ilişkili ise, bu durumda sahte (spurious) bir ilişki bulmuş oluruz. Buna **sahte(spurious) regresyon** denir.
- Regresyona bir zaman trendi ekleyerek bu sorunu aşabiliriz.

$$y_t = \beta_0 + \beta_1 x_{t1} + \beta_2 x_{t2} + \beta_3 t + u_t$$

- Eklenen zaman endeksi t'nin katsayısı y'deki x'lerle ilişkili olmayan artışı ya da azalışı verecektir.
- t' nin eklenmemesi ihmal edilmiş değişken sapmasına yol açar.

ÖRNEK 1.7: Ev yatırımları ve ev fiyatları ilişkisi,1947-88,ABD

- $invpc$: adam başına reel ev yatırımları (bin dolar)
- $price$: ev fiyatları endeksi (1982=1)
- Sahte regresyon olabilir. Trent alınca ilişki kayboldu.

Regresyona trent dahil edilmediğinde

$$\widehat{\log(\text{invpc})} = \underset{(0.43)}{-5.50} + \underset{(0.382)}{1.241}\, log(price)$$

$$n = 42 \quad R^2 = 0.208 \; \bar{R}^2 = 0.189$$

Regresyona trent eklendiğinde

$$\widehat{\log(\text{invpc})} = \underset{(0.136)}{-0.913} - \underset{(0.679)}{0.381}\, log(price) + \underset{(0.0035)}{0.0098}\, t$$

$$n = 42 \quad R^2 = 0.341 \; \bar{R}^2 = 0.307$$

Trende Sahip Değişkenlerin Regresyonda Kullanımı

- Regresyona trent eklemekle; y ve x değişkenlerinin önce trentlerini bertaraf etmek (detrending), daha sonra bu trendi alınmış değişkenler arasında regresyon tahmin etmek aynı şeydir. Aşağıdaki üç modeli düşünelim:
  1. $y_t = \alpha_0 + \alpha_1 t + u_t$
  2. $x_{t1} = \theta_0 + \theta_1 t + e_t$
  3. $x_{t2} = \gamma_0 + \gamma_1 t + h_t$
- Eğer 1. modelin tahmininde elde edilen artıklar, 2. ve 3. modelin artıklarıyla regresyona alınırsa, ulaşılan eğim katsayıları (sabit koymaya gerek yok, ancak koysak bile zaten 0 çıkar),

$$y_t = \beta_0 + \beta_1 x_{t1} + \beta_2 x_{t2} + \beta_3 t + hata$$

- Yukarıdaki modelin tahmini sonucu elde edilecek $\hat{\beta}_1$ ve $\hat{\beta}_2$ ile aynı olacaktır.

Trende Sahip Değişkenlerin Regresyonda Kullanımı

- Çoğu kez zaman serisi regresyonlarının R^2'leri kesitler-arası veriye kıyasla çok daha yüksektir.
- Bunun bir nedeni, zaman serilerinde verilerin genellikle toplulaştırılmış (aggregated) nitelikte olmasıdır.
- Toplulaştırılmış verileri açıklamak bireysel verilere göre daha kolaydır.
- Ancak R^2'yi zaman serisi regresyonlarında asıl yükselten faktör y'nin (bağımlı değişkenin) trende sahip olmasıdır.
- R^2, hata terimi varyansının y'nin varyansına göre görece büyüklüğünün bir ölçüsü idi.
- y trende sahipken $\frac{SST}{n-1}$, artık $Var(y_t)$'nin sapmasız ve tutarlı bir tahmin edicisi değildir.
- y trende sahipken, R^2'nin hesaplanmasının detayları için sf.366-367'ye bakınız.

Mevsimsellik (Seasonality)

- Belli bir zaman aralığında (çeyrek yıl, aylık, haftalık, günlük vb) gözlenmiş iktisadi veriler genellikle mevsimsellik (seasonality) izler.
- Örneğin, mevsimlere göre değişen iklim koşulları, tatillerin belli aylara toplanması (örnek, Aralık ayı için Christmas etkisi) vs. değişkenlerde sistematik mevsimsel kalıplar yaratır.
- Önemli ölçüde mevsimsel örüntü (pattern) sergileyen seriler düzeltmeye (seasonal adjustment) tabi tutulur.
- Mevsimsel düzeltme yapılmamış ham verilerle çalışılıyorsa, regresyona mevsimsel kukla değişkenler (seasonal dummies) eklemeliyiz.

Mevsimsellik (Seasonality)

- Aylık verilerin kullanıldığı bir regresyonda 11 aya ait aylık kukla değişkenler kullanırız, regresyona dahil etmediğimiz 12.ci ay (genellikle Ocak) baz kategori olur.

$$\begin{aligned} y_t &= \beta_0 + \delta_1 feb_t + \delta_2 mar_t + ... + \delta_{11} dec_t \\ &+ \beta_1 x_{t1} + ... + \beta_k x_{tk} + u_t \end{aligned}$$

- F testi ile tüm 11 kuklanın katsayılarının aynı anda sıfır olup olmadığını test ederiz. Yeterince büyük bir hesaplanan F istatistiği seride mevsimsellik olduğunu gösterir.
- Çeyrek yıllık veriler kullanılıyorsa, 3 adet mevsimsel kukla kullanacağız.
- Nasıl regresyona trent (t) eklemek değişkenlerin trentten arındırılması (**detrending**) anlamına geliyorsa mevsim kuklalarının eklenmesi de mevsimsellikten arındırma (**deseasonalizing**) demek olacaktır.