# Normallik Varsayımı ve Ençok Olabilirlik Yöntemi

## Ençok Olabilirlik Yöntemi

Ekonometri 1 – Konu 13

Sürüm 2,0 (Ekim 2011)

# UADMK Açık Lisans Bilgisi

İşbu belge, "Creative Commons Attribution-Non-Commercial ShareAlike 3.0 Unported" (CC BY-NC-SA 3.0) lisansı altında bir açık ders malzemesi olarak genel kullanıma sunulmuştur. Eserin ilk sahibinin belirtilmesi ve geçerli lisansın korunması koşulu ile özgürce kullanılabilir, çoğaltılabilir ve değiştirilebilir. Creative Commons örgütü ve "CC-BY-NC-SA" lisansı ile ilgili ayrıntılı bilgi "http://creativecommons.org" adresinde bulunmaktadır. Bu ekonometri ders notları setinin tamamına "http://www.acikders.org.tr" adresinden ulaşılabilir.

A. Talha Yalta
TOBB Ekonomi ve Teknoloji Üniversitesi
Ekim 2011 BY-NC-SA

# Ders Planı

# Ençok Olabilirlik Yaklaşımı

- İstatistikte tüm anakütleler kendilerine karşılık gelen bir olasılık dağılımı ile tanımlanırlar.
- Sıradan en küçük kareler yöntemi ise özünde olasılık dağılımları ile ilgili herhangi bir varsayım içermez.
- Bu nedenle çıkarsama yapmada SEK tek başına bir işe yaramaz.
- SEK'i genel bir tahmin süreci olarak değil de örneklem bağlanım işlevlerinin katsayılarını bulmada kullanılan bir hesaplama yöntemi olarak görmeliyiz.

# Ençok Olabilirlik Yaklaşımı

- SEK yönteminden daha güçlü kuramsal özellikler gösteren bir diğer nokta tahmincisi ise "ençok olabilirlik" (maximum likelihood), kısaca "EO" (ML) yöntemidir.
- Ençok olabilirlik yönteminin ardında yatan temel ilke şu beklentidir:

  *"Rastsal bir olayın gerçekleşmesi, o olayın gerçekleşme olasılığı en yüksek olay olmasındandır."*

- Bu yöntem 1920'li yıllarda İngiliz istatistikçi Sir Ronald A. Fisher (1890-1962) tarafından bulunmuştur.
- Ki-kare sınaması, Bayesçi yöntemler ve çeşitli ölçüt modelleri gibi birçok istatistiksel çıkarım yöntemi temelde EO yaklaşımına dayanır.

# Ençok Olabilirlik Yaklaşımı

- EO yöntemini anlayabilmek için elimizde rastsal olarak belirlenmiş bir örneklem ve dağılım katsayıları bilinen farklı anakütle adayları olduğunu varsayalım.
- Bu örneklemin farklı anakütlelerden gelme olasılığı farklı ve bazı ana kütlelerden gelme olasılığı diğerlerine göre daha yüksektir.
- Elimizdeki örneklem eğer bu anakütlelerden birinden alınmışsa, alınma olasılığı ençok olan anakütleden alınmış olduğunu tahmin etmek akılcı bir yaklaşımdır.

# Ençok Olabilirlik Yaklaşımı

Ençok olabilirlik yöntemi kısaca şöyledir:

1. Anakütlenin olasılık dağılımı belirlenir ya da bu yönde bir varsayım yapılır.
2. Eldeki örneklem verilerinin gelmiş olma olasılığının ençok olduğu anakütlenin hangi katsayılara sahip olduğu bulunur.

# İkiterimli Dağılım Örneği

Ençok olabilirlik yöntemini daha iyi anlayabilmek için şu basit örneği ele alalım:

- Elimizde, içinde siyah ya da beyaz toplam on top bulunan değişik torbalar olsun.
- Torbadaki siyah top sayısı $i = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$ olmak üzere 11 farklı torba olasıdır.
- Bu torbalardan birinden "yerine koyarak" (with replacement) dört top seçtiğimizi ve S-B-S-B sırasıyla 2 siyah top geldiğini varsayalım.
- Bu sonucun hangi torbadan gelmiş olabileceğini ençok olabilirlik yaklaşımı ile tahmin edelim.

# İkiterimli Dağılım Örneği

- Elimizdeki soru "ikiterimli" (binomial, Bi) dağılım konusudur.
- İkiterimli dağılıma göre, örnek olarak, 8'i siyah olmak üzere içinde 10 top olan bir torbadan yerine koyarak çekilen 4 toptan 2'sinin (belli bir sıra ile) siyah gelme olasılığı şudur:

$$\mathrm{Bi}(2|4, \frac{8}{10}) = \left(\frac{8}{10}\right)^2 \left(1 - \frac{8}{10}\right)^{4-2} = 0{,}0256$$

- Torbadaki siyah top oranı $p$ olsun. Örneğimizde $p$ için 11 farklı değer söz konusudur:

$$p = \{\tfrac{0}{10}, \tfrac{1}{10}, \ldots, \tfrac{10}{10}\}$$

- Bu 11 farklı torba için, 4 toptan 2'sinin siyah gelmesi durumunun gerçekleşme olasılığı şöyle gösterilebilir:

$$\mathrm{Bi}(2|4, p) = p^2(1-p)^{4-2}$$

# İkiterimli Dağılım Örneği

Çizelge: Siyah Top Sayısına Göre Olasılığın Aldığı Değerler

| Siyah Top Sayısı | Olasılık | | | | |
|---|---|---|---|---|---|
| 0 | $\mathrm{Bi}(2\|4,\frac{0}{10})$ | $=$ | $(0)^2(1)^{4-2}$ | $=$ | 0 |
| 1 | $\mathrm{Bi}(2\|4,\frac{1}{10})$ | $=$ | $(0{,}1)^2(0{,}9)^{4-2}$ | $=$ | 0,0081 |
| 2 | $\mathrm{Bi}(2\|4,\frac{2}{10})$ | $=$ | $(0{,}2)^2(0{,}8)^{4-2}$ | $=$ | 0,0256 |
| 3 | $\mathrm{Bi}(2\|4,\frac{3}{10})$ | $=$ | $(0{,}3)^2(0{,}7)^{4-2}$ | $=$ | 0,0441 |
| 4 | $\mathrm{Bi}(2\|4,\frac{4}{10})$ | $=$ | $(0{,}4)^2(0{,}6)^{4-2}$ | $=$ | 0,0576 |
| 5 | $\mathrm{Bi}(2\|4,\frac{5}{10})$ | $=$ | $(0{,}5)^2(0{,}5)^{4-2}$ | $=$ | 0,0625 |
| 6 | $\mathrm{Bi}(2\|4,\frac{6}{10})$ | $=$ | $(0{,}6)^2(0{,}4)^{4-2}$ | $=$ | 0,0576 |
| 7 | $\mathrm{Bi}(2\|4,\frac{7}{10})$ | $=$ | $(0{,}7)^2(0{,}3)^{4-2}$ | $=$ | 0,0441 |
| 8 | $\mathrm{Bi}(2\|4,\frac{8}{10})$ | $=$ | $(0{,}8)^2(0{,}2)^{4-2}$ | $=$ | 0,0256 |
| 9 | $\mathrm{Bi}(2\|4,\frac{9}{10})$ | $=$ | $(0{,}9)^2(0{,}1)^{4-2}$ | $=$ | 0,0081 |
| 10 | $\mathrm{Bi}(2\|4,\frac{10}{10})$ | $=$ | $(1)^2(0)^{4-2}$ | $=$ | 0 |

- Çizelgeye bakarak eldeki örneklemin ençok olasılıkla siyah top sayısı 5 olan torbadan alınmış olduğunu tahmin ederiz.

ULUSAL AÇIK DERS MALZEMELERİ KONSORSİYUMU

# İkiterimli Dağılım EO Tahmincisi

Tanımlamış olduğumuz ikiterimli dağılım sorusunu şimdi bir de "çözümlemesel" (analytical) olarak ele alalım:

- Elimizde içinde kaç siyah ve beyaz top olduğu bilinmeyen bir torba olsun.
- Torbadaki siyah top oranına $0 \leq p \leq 1$ diyelim.
- İlk örnekte 4 toptan oluşan bir örneklem alınmıştı. Şimdi ise örneklem büyüklüğü $n$, çıkan siyah top sayısı da $k$ olsun.
- Farklı $n$ ve $k$ sonuçları veren toplam $N$ sayıda bağımsız çekiliş yapalım.
- Ençok olabilirlik yöntemini kullanarak anakütle katsayısı $p$'yi tahmin etmek istiyor olalım.

# İkiterimli Dağılım EO Tahmincisi

- Eldeki sorunun ikiterimli dağılımı ilgilendirdiğini biliyoruz.
- İstatistikte ikiterimli dağılım, "başarı" olasılığı $p$ olan $n$ bağımsız deneyde başarılı olan $k$'lerin dağılımını gösteren bir kesikli olasılık dağılımıdır.
- Olasılık yoğunluk işlevi şudur:

$$\mathrm{Bi}(k|n,p) = \binom{n}{k} p^k (1-p)^{n-k} \tag{1}$$

- Yukarıda verilen OYİ sırasız çekilişler içindir. Matematiksel kolaylık açısından sonuçların belirli bir sırayı izlemesi gerektiğini varsayalım.
- Bu durumda kesikli OYİ şu olur:

$$p^k (1-p)^{n-k} \tag{2}$$

# İkiterimli Dağılım EO Tahmincisi

- Elimizdeki olasılık yoğunluk işlevi şuydu:
$$\mathrm{Bi}(k|n,p) = p^k(1-p)^{n-k} \tag{3}$$
- Toplam $N$ sayıdaki çekiliş için birleşik yoğunluk işlevi:
$$\mathrm{Bi}(k_i|n_i,p) = \mathrm{Bi}(k_1|n_1,p)\ \mathrm{Bi}(k_2|n_2,p)\ldots\mathrm{Bi}(k_N|n_N,p) \tag{4}$$
- Her bir $n_i$ ve $k_i$ için, (3)'ü (4)'te yerine koyalım:
$$\mathrm{Bi}(k_i|n_i,p) = p^{\sum k_i}(1-p)^{\sum n_i - k_i} \tag{5}$$
- $n_1, n_2 \ldots n_N$ ve $k_1, k_2 \ldots k_N$ değerleri veriliyken anakütle katsayısı $p$ eğer bilinmiyorsa, yukarıda gösterilen işleve "olabilirlik işlevi" (likelihood function) adı verilir:
$$\mathrm{O\dot{I}}(p) = p^{\sum k_i}(1-p)^{\sum n_i - k_i} \tag{6}$$

# İkiterimli Dağılım EO Tahmincisi

- Adından da anlaşılacağı gibi EO tahmini, verili $n_i$ ve $k_i$'leri gözleme olasılığını ençoklamaya dayanır.
- Öyleyse, hedefimiz olabilirlik işlevinin "ençoksal" (maximal) değerini bulmak olmalıdır.
- Bu da doğrudan bir türev hesabıdır.

# İkiterimli Dağılım EO Tahmincisi

- Bir işlev kendi logaritması ile “tekdüze” (monotonous) ilişkilidir. Bu nedenle olasılık işlevi yerine “log-olasılık” (log-likelihood) işlevini ençoklamak hesap kolaylığı sağlar:

$$\ln \text{Oİ}(p) = \sum_{i=1}^{N} k_i \ln(p) + \sum_{i=1}^{N} (n_i - k_i) \ln(1-p) \tag{7}$$

- (7) eşitliğinin $p$'ye göre kısmi türevini alıp sıfıra eşitleyelim:

$$\frac{\partial \ln \text{Oİ}}{\partial p} = \sum_{i=1}^{N} k_i \frac{1}{p} + \sum_{i=1}^{N} (n_i - k_i) \frac{1}{1-p}(-1) = 0 \tag{8}$$

# İkiterimli Dağılım EO Tahmincisi

- Sadeleştirmelerden sonra, EO tahmincisi $\tilde{p}$ şöyle bulunur:

$$\tilde{p} = \frac{\sum_{i=1}^{N} k_i}{\sum_{i=1}^{N} n_i} \tag{9}$$

- $p$ üzerindeki ($\sim$) "dalga" (tilde) imi, bunun bir EO tahmincisi olduğunu göstermek için kullanılmıştır.
- Görülüyor ki EO yöntemi anakütledeki siyah top oranı $k$ değerini, $N$ çekiliş sonunda bulunan siyah top sayısının çekilen toplam top sayısına oranı olarak tahmin etmektedir.

# Önümüzdeki Dersin Konusu

## Önümüzdeki ders

EO Açıklayıcı örnekler