

MIT OpenCourseWare
<http://ocw.mit.edu>

14.30 Ekonomide İstatistiksel Yöntemlere Giriş
Bahar 2009

Bu materyale atıfta bulunmak ve kullanım koşulları için <http://ocw.mit.edu/terms> sayfasını ziyaret ediniz.

14.30 Ekonomide İstatistiksel Yöntemlere Giriş

Ders Notları 17

Konrad Menzel

16 Nisan 2009

1. Merkezi Limit Teoremi

Geçen hafta binom rasgele değişkenler için DeMoivre-Laplace teoremini gördüğümüzü hatırlayınız. Bu teorem esas itibariyle, büyük n değerleri için standardize edilmiş $Y \sim B(n, p)$ dağılımlı rasgele değişkenin $Z = \frac{Y - \mathbb{E}[Y]}{\sqrt{n \text{Var}(Y)}}$ 'nin yaklaşık olarak standart normal dağılımlı olduğunu söyler. Bir binom, i.i.d. sıfır/bir rasgele değişken X_i 'lerin (başarı ile sonuçlanan "denemeler" in sayımı) toplamı olduğu için, $\frac{Y}{n}$ 'yi X_1, \dots, X_n örnekleminin ortalaması olarak düşünebiliriz.

Dolayısıyla, DeMoivre-Laplace teoremi gerçekte i.i.d. sıfır/bir rasgele değişkeninin ortalamasının bir sonucudur. Merkezi Limit Teoremi bunu sonlu varyansı olan diğer herhangi bir dağılımdan elde edilen i.i.d. sekanslarının örneklem ortalamalarına genelleştirir.

Teorem 1 (Merkezi Limit Teoremi). Varsayalım ki X_1, \dots, X_n ortalaması μ ve varyansı $\sigma^2 < \infty$ olan bir dağılımdan elde edilen n büyüklüğünde bir rasgele örneklemdir. Herhangi bir sabit x değeri için aşağıdaki yazılabilir:

$$\lim_{n \rightarrow \infty} P \left(\sqrt{n} \frac{\bar{X}_n - \mu}{\sigma} \leq x \right) = \Phi(x)$$

Burada $\sqrt{n} \bar{X}_n$ 'nin ortalaması μ ve varyansı σ^2 olan normal dağılıma yakınsayacağını söyleyebiliriz (bazıları buna "yakınsama kanunu" der). Semboller ile ifade edersek

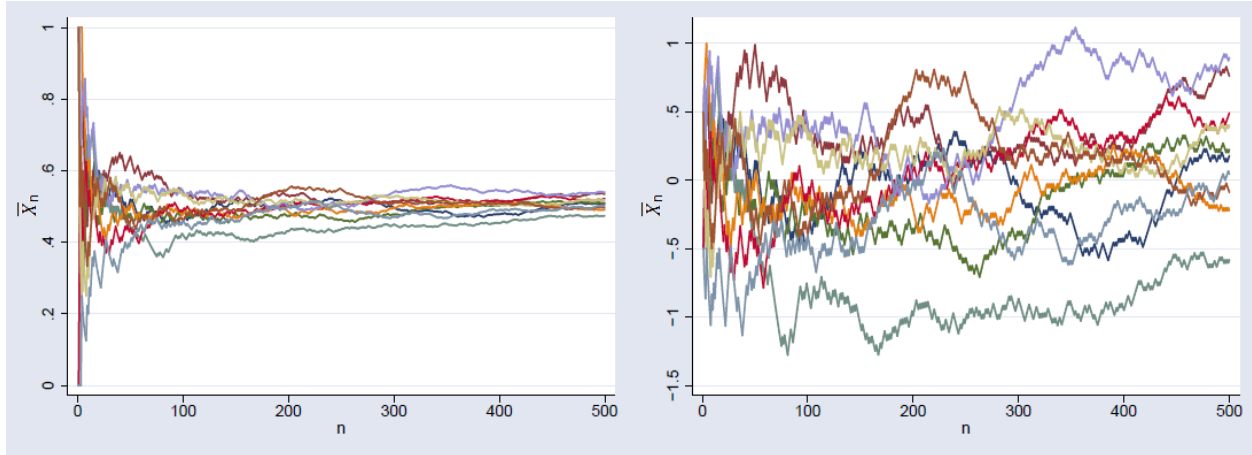
$$\sqrt{n}(\bar{X}_n - \mu) \xrightarrow{d} N(0, \sigma^2)$$

O halde bir ortalama nasıl olurda aynı anda hem sabit bir değer μ 'ya (Büyük Sayılar Kanununa göre) ve hem de varyansı 1 olan (merkezi limit teoremine göre) rasgele bir değişkene yakınsar? Buradaki önemli detay merkezi limit teoremi ile ilgilidir yani dağılımın bir noktaya düşmesini (Büyük Sayılar Kanununda bu olur) veya sonsuza

kadar patlamasını engelleyecek en uygun oranı verecek şekilde örneklem ortalamasını \sqrt{n} 'nin çarpımı kadar yukarı çekeriz.

Bir dağılımın ortalamasını sınırlandırmak için niye normal dağılım başlamak için en uygun adaydır? İki bağımsız normal rasgele değişkenin toplamının yine normal bir dağılımı olduğunu (her ne kadar varyansları farklı olsa da, sadece standardize edilmiş ortalamaya baktığımız için bu o kadar da önemli değildir) tartıştığımızı tekrar hatırlayınız. Yani bükülme (yani dağılımdan bağımsız rasgele değişken ekleme) açısından normal dağılım ailesi istikrarlıdır. Bunun diğer dağılımların çoğu için doğru olmadığını belirtelim (uniform veya üstel gibi).

Örneklem ortalaması tek tek gözlemlerin ağırlıklandırılmış ortalaması olduğu için, diyelim ki örnekleme n 'den $2n$ 'e çıkarırsak, X_{n+1}, \dots, X_{2n} sekansının ortalaması birinci ortalamaya eklenir ve 2'ye bölünür. Dolayısıyla, eğer büyük n 'ler için bile, \bar{X}_n 'in dağılımının elde edilen iki bağımsız değişkenin toplamının aynı dağılım ailesinde olmayan türden olduğunu varsaymış olsaydık, örneklem ortalamasının dağılımı yine de ihtiyari büyük n değerleri için çok değişirdi ve dolayısıyla istikrarlı bir limite doğru gitmeyebilirdi. Bu durum, ortalamasının dağılımının limite normal dağılıma yaklaşmasının niye kabul edilebilir olduğu konusunda yeterince bilgi veriyor olması gerekir .



Şekil 1. n para atışının tura sayısı: örneklem ortalaması \bar{X}_n (solda) ve standardize edilmiş örneklemin ortalaması $\sqrt{n}\bar{X}_n$ (sağda)

Örnek 1. Varsayalım ki X_1, \dots, X_n i.i.d. olan rasgele değişkenlerdir ve $X_i \sim U[0, 1]$ dağılımı uniformdur, dolayısıyla p.d.f şöyle olur:

$$f_X(x) = \begin{cases} 1 & \text{eğer } 0 \leq x \leq 1 \text{ ise} \\ 0 & \text{diğer bütün durumlarda} \end{cases}$$

Şimdi kısmi toplamların p.d.f.sini hesaplamak için 10'nuncu dersteki bükülme formülünü kullanabiliriz,

$$S_k = X_1 + X_2 + \dots + X_k$$

$k = 2$ için, aşağıdakini elde ederiz (integralin limitlerine dikkat etmek gerekir)

$$\begin{aligned} f_{S_2}(s_2) &= \int_{-\infty}^{\infty} f_X(s_2 - w)f_X(w)dw = \int_{\max\{s_2-1,0\}}^{\min\{s_2,1\}} 1dw \\ &= \min\{s_2, 1\} - \max\{s_2 - 1, 0\} = \begin{cases} s_2 & \text{eğer } 0 \leq s_2 \leq 1 \text{ ise} \\ 2 - s_2 & \text{eğer } 1 \leq s_2 \leq 2 \text{ ise} \\ 0 & \text{diğer bütün durumlarda} \end{cases} \end{aligned}$$

Sonraki hesaplamalar çok karmaşık olabilir çünkü integralin limitlerini ve yoğunluktaki kırılma noktalarını sürekli kontrol etmek gerekir. Bazı hesaplamalardan sonra, $k = 3$ için şunu elde ettik,

$$f_{S_3}(s_3) = \int_{-\infty}^{\infty} f_{S_2}(s_3 - w)f_X(w)dw = \begin{cases} \frac{s_3^2}{2} & \text{eğer } 0 \leq s_3 \leq 1 \text{ ise} \\ -\frac{3}{2} + 3s_3 - \frac{s_3^2}{2} & \text{eğer } 1 \leq s_3 \leq 2 \text{ ise} \\ \frac{9}{2} - 3s_3 + \frac{1}{2}s_3^2 & \text{eğer } 2 \leq s_3 \leq 3 \text{ ise} \\ 0 & \text{diğer bütün durumlarda} \end{cases}$$

Rasgele değişkenlerin toplamlarının beklenen değeri ile ilgili kurala göre

$$\mathbb{E}[S_k] = k\mathbb{E}[X_1] = \frac{k}{2}$$

Ayrıca, X_1, X_2, \dots, X_k bağımsız oldukları için, toplamın varyansı ile ilgili kuralı kullanabiliriz

$$\begin{aligned} \text{Var}(S_k) &= \text{Var}(X_1 + X_2 + \dots + X_k) = \text{Var}(X_1) + \text{Var}(X_2) + \dots + \text{Var}(X_k) \\ &= k\text{Var}(X_1) = k \int_0^1 \left(t - \frac{1}{2}\right)^2 dt = k \left(\frac{1}{3} - \frac{1}{2} + \frac{1}{4}\right) = \frac{k}{12} \end{aligned}$$

Bundan ötürü, S_k 'nin standardize edilmiş hali olan Z_k aşağıdaki ile verilir:

$$Z_k = \frac{S_k - \frac{k}{2}}{\sqrt{\frac{k}{12}}}$$

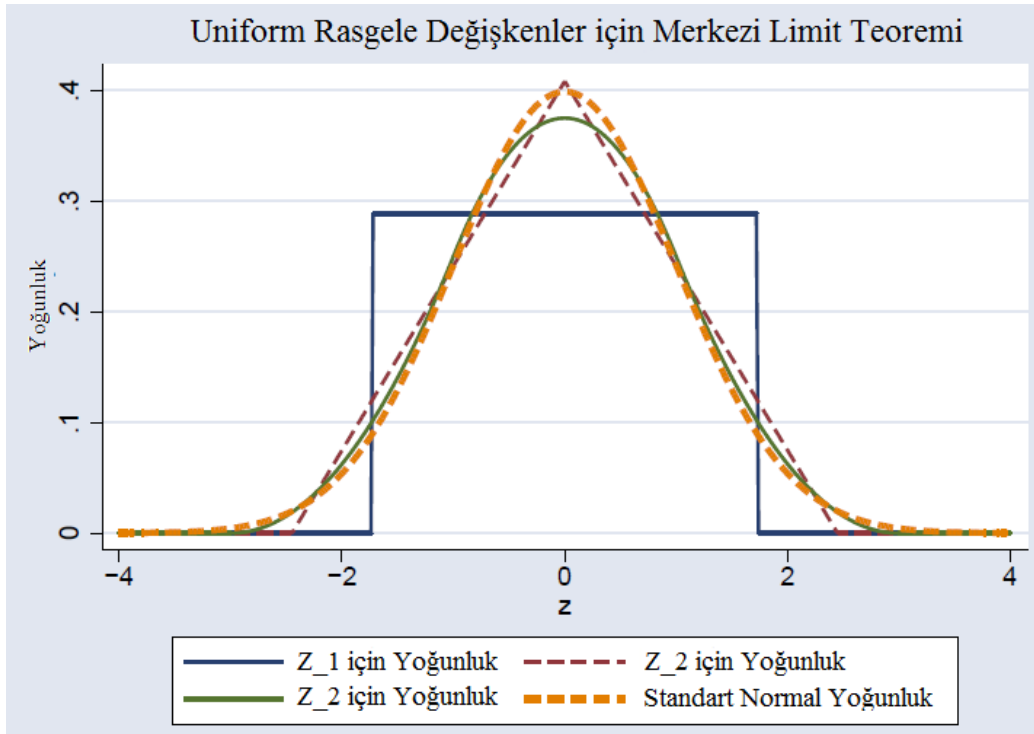
Buradan hareketle, deęişken deęiřtirme formülünü kullanarak standardize edilmiş Z_1 , Z_2 , Z_3 'ün yoğunluklarını hesaplayabiliriz (türevin sadece $\sqrt{12/k}$ 'ya eşit olduğuna dikkat ediniz):

$$f_{Z_1}(z) = \begin{cases} \frac{1}{\sqrt{12}} & \text{eđer } -\sqrt{3} \leq z \leq \sqrt{3} \text{ ise} \\ 0 & \text{diđer bütün durumlarda} \end{cases}$$

$$f_{Z_2}(z) = \begin{cases} \frac{1}{\sqrt{6}} + \frac{z}{6} & \text{eđer } -\sqrt{6} \leq z \leq 0 \text{ ise} \\ \frac{1}{\sqrt{6}} - \frac{z}{6} & \text{eđer } 0 \leq z \leq \sqrt{6} \text{ ise} \\ 0 & \text{diđer bütün durumlarda} \end{cases}$$

$$f_{Z_3}(z) = \begin{cases} \frac{1}{4} \left(\frac{3}{2} + \frac{z}{2} \right)^2 & \text{eđer } -3 \leq z \leq -1 \text{ ise} \\ \frac{3}{8} - \frac{z^2}{8} & \text{eđer } -1 \leq z \leq 1 \text{ ise} \\ \frac{1}{2} \left(\frac{9}{8} - \frac{3}{4}z + \frac{z^2}{8} \right) & \text{eđer } 1 \leq z \leq 3 \text{ ise} \\ 0 & \text{diđer bütün durumlarda} \end{cases}$$

Bunun grafiksel olarak nasıl görüldüğüne bakalım:



Uniform rasgele deęişkenlerin standardize edilmiş toplamlarının p.d.f.si sadece 3 kadar bağımsız çekilişin toplamlarının standart normal p.d.f.sine çok benzemektedir. Bu bir

sürprizdir çünkü uniform yoğunluğun kendisi hiçbir şekilde normal rasgele değişkeninkine benzemez.

Son örnek, 3 kadar az (en azından optik olarak) sayıdaki n için normal tahminin oldukça iyi görünmesi konusunda biraz aldatıcı olsa da, $n \rightarrow \infty$ durumunda, genellikle $n \approx 40$ veya daha yüksek değerler için tahminin doğru olduğunu ifade ederiz.

Özetleyecek olursak, örneklemin gerçek ortalamasının p.d.f.sini hesaplamak istemediğimiz zaman Merkezi Limit Teoremi özellikle yararlıdır. Bunun gerçekleşebileceği iki durum vardır

- Gerçek p.d.f.'yi *hesaplamayabiliriz* çünkü X 'lerin tam dağılımını bilmiyoruzdur,
- Gerçek p.d.f.yi hesaplamak *istemeyebiliriz* çünkü hesaplamalar çok karmaşık olabilir- Bu genel bükülme formülü (örneğe bakınız), ve diğer birçok kesikli örnek (son derste ki Binom örnek gibi) için nerdeyse tartışmasız doğrudur.

2. Tahmin

Bu derste şimdiye kadar, bir rasgele değişkenin dağılımının parametrelerini bildiğimizi varsayarak başladık. Örneğin $X \sim P[\lambda]$ olduğunu biliyorduk ve daha sonra o dağılımdan olasılıkları ve rasgele değişkenin diğer özelliklerin hesaplıyorduk. Şimdi ise probleme tersten bakacağız:

θ parametreleri bilinmeyen bir dağılımdan elde edilen bir i.i.d. örnekleme sahip olduğumuzu varsayacak olursak, işlediğimiz dağılım ailelerinde yer alan hangi θ değerinin veriyi üretmiş olabileceği konusunda nasıl "makul" bir cevap elde ederiz?

Örnek 2. *Eğer bir madeni para verilmişken tek bir atışın tura gelme olasılığını bilmiyorsak, çok sayıda atış yapabiliriz. O zaman turların oranını, $\hat{p} = \frac{\text{Tura Sayısı}}{\text{Atış Sayısı}}$, daha sonra açıklayacağımız mantıkla $P(\text{Tura})$ olasılığı için "iyi bir tahmin" olabileceğini düşünebiliriz.*

$f(x|\theta)$ p.d.f.leri veri iken, bir *parametre* bir dağılımlar ailesinin sabit endeksidir. Burada parametreleri genellikle $\theta_1, \dots, \theta_k$ olarak ifade ederiz.

Örnek 3.

- *Binom dağılım için parametreler n deneme sayısı ve p başarı oranıdır.*

$$f_X(x|n, p) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x} & \text{eğer } x = 0, 1, \dots, n \text{ ise} \\ 0 & \text{diğer bütün durumlarda} \end{cases}$$

- Normal dağılım için parametreler ortalama μ ve standart sapma σ 'dır.

$$f_X(x|\mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

- Poisson dağılımı bir tek parametreye sahiptir, λ

$$f_X(x) = \begin{cases} \frac{\lambda^x e^{-\lambda}}{x!} & \text{eğer } x \in \{0, 1, 2, \dots\} \text{ ise} \\ 0 & \text{diğer bütün durumlarda} \end{cases}$$

İstatistiğin çoğu, bilinen dağılım ailelerinin hangi üyesinin gözlemlenmiş bir sürecin veya olayın doğru olasılık dağılımını vereceğini belirlemekle ilgilidir. Sembolik olarak, $X \sim f[x|\theta_0]$ iken biz parametre değeri θ_0 'ı bulmak isteriz. Bu durum, "dağılımı karakterize eden parametreleri tahmin etme" problemidir.

Her zaman bir rasgele örneklem X_1, \dots, X_n ile başlayacağız ve her zaman aşağıdaki varsayımda bulunacağız:

$$\text{bilinmeyen } \theta_0 \in \theta \text{ için } X \sim f[x|\theta_0]$$

Tanım 1. θ 'nin tahmin edicisi $\hat{\theta}$ bir istatistiktir (yani X_1, \dots, X_n 'in bir fonksiyonudur),

$$\hat{\theta} = \hat{\theta}(X_1, \dots, X_n)$$

Bir örnekleme bir tahmin edicinin *gerçekleşmesi*, $\hat{\theta}(x_1, \dots, x_n)$, θ 'nın *tahmini* olarak adlandırılır.

Rasgele bir örneklemin fonksiyonu olarak, tahmin edici uygun bir rasgele değişkendir, bu nedenle genel olarak p.d.f.si cinsinden dağılımını açıklamak ve dağılımının momenti ile ilgileneceğimize dikkat ediniz.

Örnek 4. Varsayalım ki $X_i \sim \text{Bernoulli}(\theta_0)$ 'dir, yani X sıfır/bir rasgele değişkendir ve 1 değerini θ olasılıkla alır ve p.d.f.si şöyledir:

$$f_X(x) = \begin{cases} \theta_0 & \text{eğer } x = 1 \text{ ise} \\ 1 - \theta_0 & \text{eğer } x = 0 \text{ ise} \\ 0 & \text{diğer bütün durumlarda} \end{cases}$$

θ_0 'ını nasıl tahmin ederiz?

Örnekleme ortalamasını kullanabiliriz,

$$\hat{\theta}(X_1, \dots, X_n) = \frac{1}{n} \sum_{i=1}^n X_i$$

Örneğin 5 Bernoulli denemesi 1, 0, 0, 1, 1 için şunu buluruz,

$$\hat{\theta}(1, 0, 0, 1, 1) = \frac{3}{5}$$

5 gözlemlenmiş bir örneklemede $\hat{\theta}$ tahmin edicisi bir rasgele değişken olduğu için, p.d.f.sini elde edebiliriz: hatırlayınız $S_5 \equiv (\sum_{i=1}^5 X_i)$ için $S_5 \sim B(5, \theta_0)$ 'dır. Kesikli rasgele değişkenlerin p.d.f.sini bulmanın yöntemlerini $\hat{\theta} = (S_5)/5$ 'e uygularsak, aşağıdakini elde ederiz:

$$f_{\hat{\theta}}(t) = \begin{cases} \binom{5}{5t} \theta_0^{5t} (1 - \theta_0)^{5(1-t)} & \text{eğer } t \in \{0, \frac{1}{5}, \frac{2}{5}, \frac{3}{5}, \frac{4}{5}, 1\} \text{ ise} \\ 0 & \text{diğer bütün durumlarda} \end{cases}$$

Özellikle, tahmin edicinin dağılımı gerçek θ_0 olasılığına bağlıdır – o olasılık $[0,1]$ aralığında herhangi bir yerde olabilir- fakat sadece 6 farklı kesikli değer alabilir.

Örnek 5. Eğer $X \sim U[0, \theta]$ dağılımı parametreye bağlı olarak bir aralıkta uniform ise, p.d.f.'si şöyledir:

$$f_X(x) = \begin{cases} \frac{1}{\theta} & \text{eğer } 0 \leq x \leq \theta \text{ ise} \\ 0 & \text{diğer bütün durumlarda} \end{cases}$$

θ 'yi nasıl tahmin edebiliriz? Aşağıdakileri kullanabiliriz,

$$\begin{aligned} \hat{\theta}_1 &= \max\{X_1, \dots, X_n\} \\ \hat{\theta}_2 &= 2\bar{X}_n \end{aligned}$$

Diyelim ki, dağılımdan üç örnek seçtik, 0.2, 0.6, 0.4. O zaman $\hat{\theta}_1 = 0.6$ ve $\hat{\theta}_2 = 0.8$ olurdu ve böylece aynı parametre için iki tahmin edici farklı cevap verirdi. Bu farklı tahmin ediciler arasından nasıl seçim yapacağız? Buna biraz sonra geri döneceğiz.

- $\hat{\theta}(X_1, \dots, X_n)$ fonksiyonlarına nasıl ulaşırsınız?
- Bu tahmin edicilerin makul olduğunu nasıl belirleyebiliriz?

- Aynı parametre için bir veya iki tahmin edici arasından nasıl seçim yapmalıyız?

3. Tahmin Edicilerin Genel Özellikleri

X 'in beklenen değerini *parametre* θ ile ifade edeceğiz, yani eğer gerçek parametre θ 'ya eşitse X 'in beklenen değeri şöyledir:

$$\mathbb{E}_\theta[X] = \int_{-\infty}^{\infty} x f_X(x|\theta) dx$$

Aynı şekilde, parametre θ ile varyansı şöyle yazacağım:

$$\text{Var}_\theta(X) = \int_{-\infty}^{\infty} (x - \mathbb{E}_\theta[X])^2 f_X(x|\theta) dx$$

Bir tahmin edicinin *sapması*(yanlılığı Ç.N.) onun beklenen değeri ile gerçek değeri arasındaki farktır.

$$\text{Sapma}(\hat{\theta}) = \mathbb{E}_{\theta_0}[\hat{\theta}] - \theta_0$$

Elbette ki, ortalama olarak parametreyi doğru veren bir tahmin ediciyi isteriz, yani ideal olanı sapmanın sıfır olmasıdır.

Tanım 2. Bir tahmin edici $\hat{\theta} = \hat{\theta}(X_1, \dots, X_n)$ θ için *sapmasızdır* eğer bütün θ_0 değerleri için

$$\mathbb{E}_{\theta_0}[\hat{\theta}] = \theta_0$$

Örnek 6. Varsayalım ki, X_1, \dots, X_n $N(\mu, \sigma^2)$ dağılımından elde edilen bir i.i.d. örneklemdir. Geçen hafta örneklem ortalamasının beklenen değerinin herhangi bir μ değeri için

$$\mathbb{E}_{\mu, \sigma^2}[\bar{X}_n] = \mathbb{E}_{\mu, \sigma^2}[X] = \mu$$

olduğunu görmüştük, bu nedenle \bar{X}_n normal dağılımın ortalaması μ için *sapmasız bir tahmin edicidir*.

Örnek 7. İ.i.d. olan X_1, \dots, X_n rasgele örneklemeden μ ortalaması bilinmeyen $X \sim N(\mu, \sigma^2)$ dağılımı için varyans parametresi σ^2 'yi tahmin etmek istediğimizi varsayalım. $\sigma^2 = \mathbb{E}[(X - \mathbb{E}[X])^2]$ olduğu için, sezgisel olarak ilgi çekici bir tahmin edici aşağıdaki gibi olabilir

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$$

(burada örneklem ortalamasını gerçek beklenen değeri ile yer değiştirttik). Eğer dağılımın gerçek parametreleri (μ_0, σ_0^2) ise, bu tahmin edicinin beklenen değeri nedir?

$\mathbb{E}[X^2] = \mathbb{E}[X]^2 + \text{Var}(X)$ ilişkisini hatırlayınız, böylece

$$\begin{aligned} \mathbb{E}[\hat{\sigma}^2] &= \mathbb{E} \left[\frac{1}{n} \sum_{i=1}^n (X_i^2 - 2X_i\bar{X} + \bar{X}^2) \right] \\ &= \frac{1}{n} \sum_{i=1}^n \mathbb{E}[X_i^2 - \bar{X}^2] \\ &= \frac{1}{n} \sum_{i=1}^n \mathbb{E}[X_i^2] - \mathbb{E}[\bar{X}^2] \\ &= \frac{1}{n} \sum_{i=1}^n (\mu^2 + \sigma^2) - \left(\mu^2 + \frac{1}{n} \sigma^2 \right) \\ &= \mu^2 + \sigma^2 - \mu^2 - \frac{\sigma^2}{n} = \frac{n-1}{n} \sigma^2 \end{aligned}$$

Dolayısıyla $\hat{\sigma}^2$ σ^2 için sapmasız bir tahmin edici değildir, fakat $\hat{\sigma}^2$ için çok kolay bir sapmasız tahmin edici oluşturabiliriz.

$$\tilde{\sigma}^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2$$

Bu sapma nereden geliyor? Genel konuşmak gerekirse, bunun nedeni kare parantezin içindeki μ 'yü "gürültülü" bir tahmin olan $\hat{\mu} = \bar{X}_n$ ile yer değiştirmemizdir. Eğer μ_0 bilinseydi, tahmin edici $\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \mu_0)^2$ 'nin σ için sapmasız olacağını kendiniz kontrol edebilirsiniz.

Ortalamayı tahmin etmek veriden bir "serbestlik derecesi" götürür- örneğin eğer elimizde sadece tek gözlemlili bir örneklem olsaydı, tahmin edilen ortalama o gözleme eşit olurdu

ve varyansın “saf” tahmin edicisi bize $\hat{\sigma}^2 = 0$ 'ı verirdi, açıkçası bu da doğru cevap değildir.

Sapmasızlık bizim ilgilendiğimiz tek şey olmayabilir. Ortalama olarak tahmin edicinin gerçek parametreye eşit olması belli bir örneklem içinde öyle olacağı anlamına gelemeyeceğinden, gerçekte tahmin gerçek parametreye yakın olacaktır.

Tanım 3. Eğer n 'yi artırırken tahmin edici olasılık bakımından θ_0 'a yakınsarsa, bir X_1, \dots, X_n örneklemini için, $\hat{\theta}$ 'nin θ için tutarlı bir tahmin edici olduğunu söyleyebiliriz. Bütün $\varepsilon > 0$ değerleri için

$$\lim_{n \rightarrow \infty} P_{\theta_0} \left(|\hat{\theta}(X_1, \dots, X_n) - \theta_0| < \varepsilon \right) = 1$$

Kelimeler ile ifade edecek olursak, yeterince büyük bir örneklemde, tutarlı bir tahmin edici yüksek bir olasılıkla gerçek parametreye çok yakın bir aralığın içinde olacaktır. Dikkat edilecek olursa sapmasızlık ve tutarlılık iki çok farklı örtüşen kavramdır, ancak biri diğerini ima etmez.

Örnek 8. Uniform dağılım için söz konusu olan tahmin edicilerimizden birine geri dönelim, $X \sim U[0, \theta_0]$. Eğer aşağıdaki ifadeye bakarsak,

$$\hat{\theta}_1 = \max\{X_1, \dots, X_n\}$$

$\hat{\theta}_1$ 'in θ için sapmasız olmadığını kolaylıkla görebiliriz, çünkü uniform dağılımın doğasından ötürü, X_i 'nin bütün olası değerleri θ_0 'ten küçüktür. Bu nedenle, n ne kadar büyük olursa olsun, $P(\max\{X_1, \dots, X_n\} < \theta_0) = 1$ 'dir. Böylece Beklenen değer $\mathbb{E}_{\theta_0}[\hat{\theta}_1] < \theta_0$ 'dir. Ancak, $\hat{\theta}_1$ θ_0 için tutarlıdır: Uniform dağılımından elde edilen bir tek gözlem olan X için c.d.f.'nin $F_X(x) = x/\theta_0$ olduğunu kolayca görebiliriz. $Y_n := \max\{X_1, \dots, X_n\}$ örneklemin n 'nci sıra istatistiği olduğu için, daha önceki tartışmalarımızdan $0 \leq y \leq 1$ için $F_{Y_n}(y) = (F_X(y))^n = (y/\theta_0)^n$ 'i elde ederiz. 1 olasılıkla $\hat{\theta}_1 < \theta_0$ olduğu için, herhangi bir örneklem sayısı n ve herhangi bir $\varepsilon > 0$ için aşağıdakini hesaplayabiliriz,

$$P(|\hat{\theta}_1 - \theta_0| > \varepsilon) = P(Y_n < \theta_0 - \varepsilon) = \left(\frac{\theta_0 - \varepsilon}{\theta_0} \right)^n \equiv p^n$$

burada $\varepsilon > 0$ olduğu için $p := \frac{\theta_0 - \varepsilon}{\theta_0} < 1$ 'dir. Bu nedenle, n 'i artırırken, θ_0 'dan ε 'dan daha fazla sapmanın olasılığı yok olur, ve dolayısıyla $\hat{\theta}_1$ tutarlıdır.

Örnek 9. Büyük Sayılar Kanununa göre, örneklem ortalaması olasılık olarak $\mathbb{E}[X] = \mu$ 'ya yakınsar. Dolayısıyla, $N(\mu, \sigma^2)$ olan rasgele değişkeninin bir i.i.d. olan X_1, \dots, X_n örneklemini için, örneklem ortalaması μ 'nün tutarlı bir tahmin edicisidir.

Ayrı bir seçenek olarak, “makul olmayan” bir $\hat{\theta}(X_1, \dots, X_n)$ tahmin edicisine bakalım. Bu durumda,

$$\mathbb{E}[\tilde{\theta}(X_1, \dots, X_n)] = \mathbb{E}[X_n] = \mu$$

Dolayısıyla bu tahmin edici sapmasızdır. Ancak, herhangi bir n örneklem büyüklüğü için, tahmin edicinin dağılımı söz konusu $X \sim N(\mu, \sigma)$ 'in aynısıdır, bu nedenle örneğin $\varepsilon = \sigma_0$ için, bütün n 'lerin olasılığı

$$\begin{aligned} P(|\tilde{\theta}(X_1, \dots, X_n) - \mu_0| < \sigma_0) &= P(\mu_0 - \sigma_0 \leq X_n \leq \mu_0 + \sigma_0) \\ &= P\left(-1 < \frac{X_n - \mu_0}{\sigma_0} < 1\right) = P(-1 < Z < 1) \\ &= \Phi(1) - \Phi(-1) \approx 0.6825 \ll 1 \end{aligned}$$

burada standardize edilen $Z := \frac{X_n - \mu_0}{\sigma_0} \sim N(0, 1)$ dağılımlıdır. Bu argümanlardan hareketle, $\tilde{\theta}$ sapmasızdır ancak tutarlı değildir.