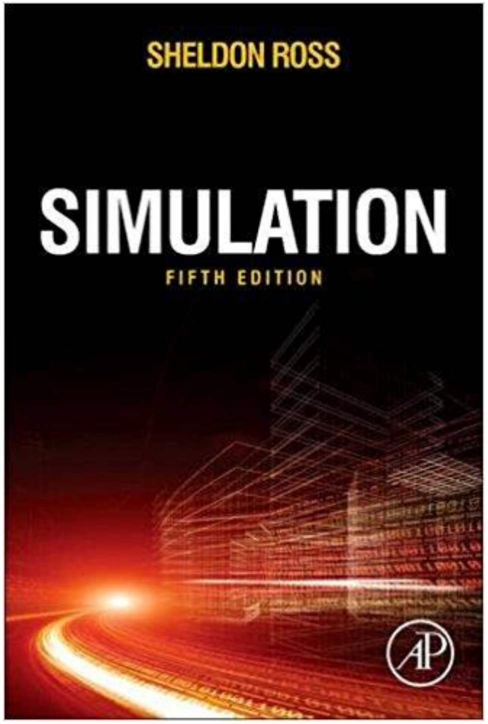
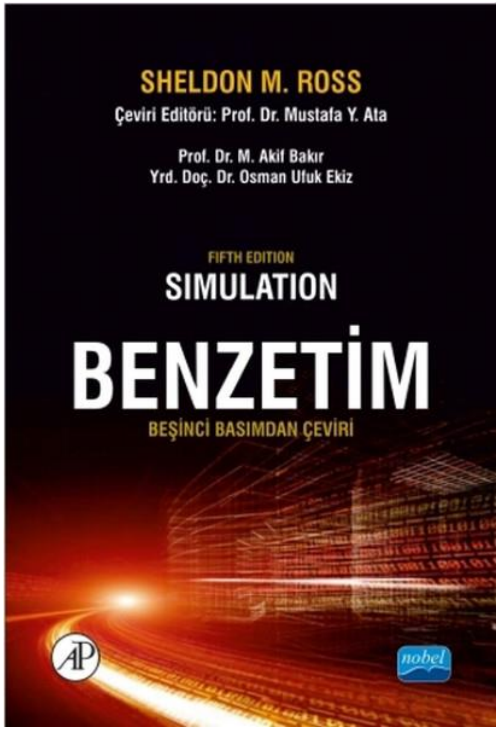
SAB311 – BENZETİM TEKNİKLERİ

Prof.Dr. Fatih TANK
Ankara Üniversitesi
Uygulamalı Bilimler Fakültesi

Bu ders notları hazırlanırken
«Simulation (S.Ross)»
kitabının çevirisi olan
«Benzetim (M.Y.Ata, M.A.Bakır, O.U.Ekiz)»
kitabından yararlanılmıştır.





Bölüm 8
BENZETİM TEKNKLERİ İLE ÜRETİLEN VERİLERİN İSTATİSTİKSEL ANALİZİ

Giriş

Bu bölümde benzetimin ne zaman durdurulacağı—yani, uygun bir k değerine karar verme sorununu—ele alacağız. Ne zaman durduracağımız kararına yardımcı olacak, θ tahmin edicimizin niteliğini ele almalıyız. Ayrıca belirli bir güven düzeyinde öngörebileceğimiz θ'yı içeren bir aralığı da nasıl bulacağımızı göstereceğiz.

8.1 Örnek Ortalaması ve Örnek Değişkesi

$D[X_i]$ —olsun. n tane verinin aritmetik ortalaması olan

$$\overline{X} \equiv \sum_{i=1}^{n} \frac{X_i}{n}$$

değerine, *örnek ortalaması* denir.

Tanım

$$S^2 = \frac{\sum_{i=1}^{n} (X_i - \overline{X})^2}{n-1}$$

şeklinde tanımlanan S^2'ye örnek değişkesi denir.

Şimdi, ispatı okuyucuya alıştırma olarak bırakılan,

$$\sum_{i=1}^{n} (X_i - \overline{X})^2 = \sum_{i=1}^{n} X_i^2 - n\overline{X}^2 \tag{8.4}$$

matematiksel özdeşliğini kullanılarak örnek değişkesinin σ^2'nin sapmasız bir tahmin edicisi olduğunu gösterelim.

Yeni Veri Üretmeyi Durdurma Zamanını Belirlemek için Bir Yöntem

1. Tahmin edicinin standart sapması için kabul edilebilir bir d değeri seçilir.
2. En az 100 veri değeri üretilir.
3. k tane değere dayalı örnek standart sapması S olmak üzere, $S/\sqrt{k} < d$ oluncaya kadar yeni veri üretmeye devam edilir.

$\overline{X} = \sum_{i=1}^{k} X_i/k$ θ'nın tahminidir.

Örnek 8a Saat 17:00'den sonra hiç bir yeni müşterinin girmesine izin verilmeyen bir hizmet dizgesi olsun. Aynı olasılık yasasının her gün geçerli olduğu ve dizgeden ayrılacak son müşterinin beklenen ayrılma zamanını tahmin etmek isteyelim. Ayrıca, tahminimizin gerçek değerden 15 saniyeden daha fazla olmayacağından en az %95 emin olmak istediğimizi varsayalım.

Yukarıdakinin gereğini yerine getirmek için, k en az 100 ve S, (saniye cinsinden ölçülen) k tane verinin örnek ortalaması olmak üzere, $1.96S/\sqrt{k} < 15$ olana kadar, son müşterinin ayrıldığı zamana ilişkin veriyi (her seferinde bir benzetim gerçekleştirerek toplam k kere) sürekli üretmemiz gerekir. □

Örnek 8b Eğer ilk üç verinin değerleri $X_1 = 5$, $X_2 = 14$, $X_3 = 9$, ise o zaman, eşitlik (8.6) ve (8.7)

$$\overline{X}_1 = 5$$

$$\overline{X}_2 = 5 + \frac{9}{2} = \frac{19}{2}$$

$$S_2^2 = 2\left(\frac{19}{2} - 5\right)^2 = \frac{81}{2}$$

$$\overline{X}_3 = \frac{19}{2} + \frac{1}{3}\left(9 - \frac{19}{2}\right) = \frac{28}{3}$$

$$S_3^2 = \frac{81}{4} + 3\left(\frac{28}{3} - \frac{19}{2}\right)^2 = \frac{61}{3}$$

sonucunu verir. □

Örnek 8c Örnek 8a'daki gibi, saat 17:30'da mağazada bir müşterinin bulunma olasılığını tahmin etmek istediğimizi varsayalım. Bunun için,

$$X_i = \begin{cases} 1, \text{ inci gün saat 17:30'da bir müşteri bulunuyorsa} \\ 0, \text{ aksi durumda} \end{cases}$$

biçiminde tanımlanmak üzere, ardışık günlerin benzetimini yapmamız gerekir. $o_k = \overline{X}_k$, saat 17:30'da mağazada bir müşterinin bulunduğu günlerin k gün içindeki oranı ve d de o_k tahmin edicisinin standart sapmasının kabul edilebilir bir değeri olmak üzere, en az 100 günün benzetimi yapılmalı ve $[o_k(1 - o_k)]^{1/2} < d$ oluncaya kadar benzetime devam edilmelidir. □

8.2 Yığın Ortalaması için Aralık Tahminleri

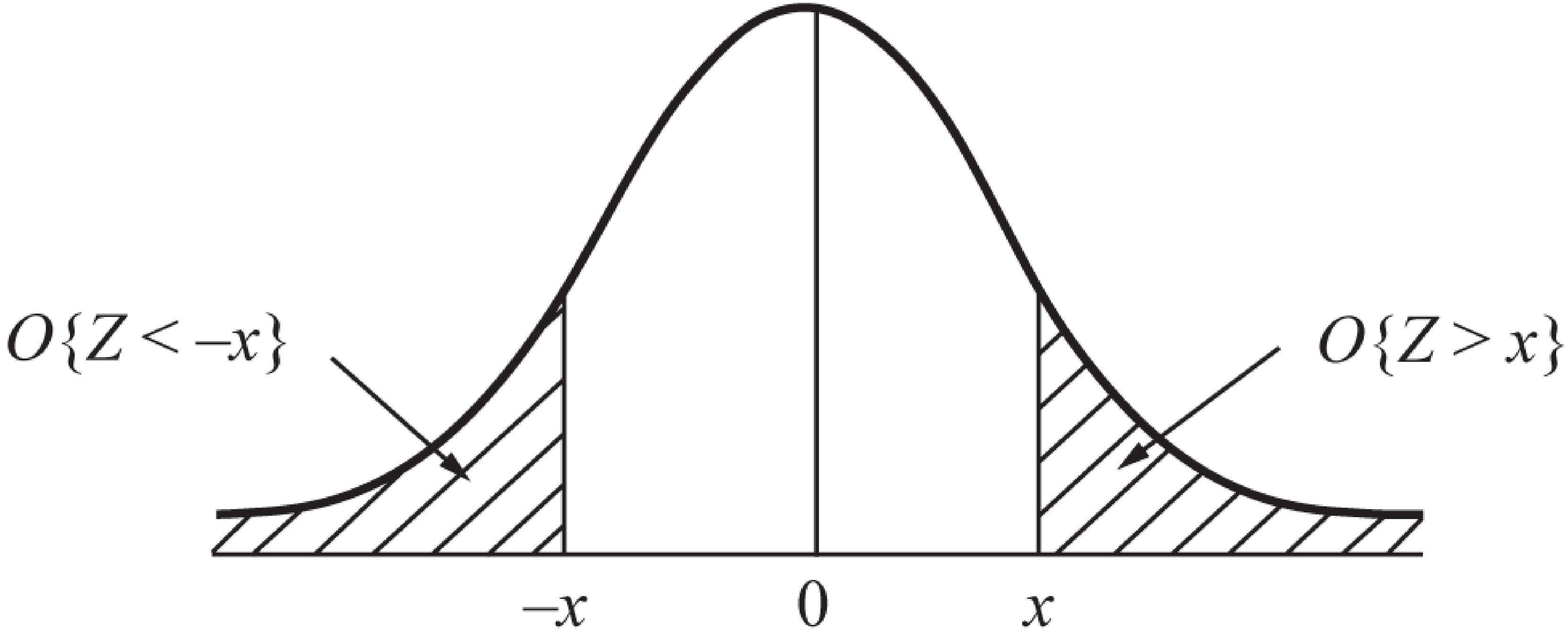
Böyle bir aralığı bulabilmek için tahmin edici $\overline{X}$'nın dağılımını (yaklaşık da olsa) bulmamız gerekir. Bunu belirleyebilmek için öncelikle Eşitlik (8.1) ve (8.2)'den

$$B[\overline{X}] = \theta,\ D[\overline{X}] = \frac{\sigma^2}{n}$$

ve böylece merkez erey teoreminden, n'in büyük değerleri için

$$\sqrt{n}\frac{(X - \theta)}{\sigma} \dot{\sim} N(0, 1)$$

olduğu anımsanmalıdır.



Şekil 8.1. Standart normal yoğunluk.

Tanım *Eğer* $\overline{X} = \overline{x}$ *ve* $S = s$ *örnek ortalaması ve örnek standart sapmasının gözlenen değerleri ise* $\overline{x} \pm z_{\alpha/2} s/\sqrt{n}$ *aralığına* θ*'nın (yaklaşık) yüzde* $100(1-\alpha)$*'lık güven aralığı denir.*

Teknik Bir Uyarı Örnek çapının üretilen verilerin değerlerine açıkça bağlı rasgele bir değişken olduğu yukarıdaki durumda kuramı örnek çapının sabitliği varsayımına dayalı yaklaşık bir güven aralığı kullanmamızdan, istatistik bilgisi daha ileri olan
okuyucu kuşku duyabilir. Ne var ki, örnek çapının büyük olması durumunda bu durum haklı görülebilir ve benzetim bakış açısından bu ince ayrıntıyı güvenle göz ardı edebiliriz

8.3 Hata Kareler Ortalamasının Tahmininde Özcül Yöntem

F_e deneysel dağılım işlevini

$$F_e(x) = \frac{i\colon X_i \leqslant x \text{ koşulunun sağlandığı gözlemlerin sayısı}}{n}$$

ile tahmin edebiliriz.

Örnek 8d $\theta(F) = B[X]$'yi örnek ortalaması $\overline{X} = \sum_{i=1}^{n} X_i/n$'yi kullanarak tahmin etmek isteyelim. Eğer x_i, $i = 1, \ldots, n$'ler gözlenmiş verilerse, o zaman deneysel dağılım F_e, $x_1, \ldots, x_n$ örneğindeki her noktaya (x_i'lerin hepsi farklı değilse aynı olanların ağırlıklarını birleştirerek) $1/n$ ağırlığını verir. Dolayısıyla, $\theta(F_e) = \overline{x} = \sum_{i=1}^{n} x_i/n$, F_e'nin ortalaması olur ve böylece hata kare ortalamasının özcül tahmini OHK(F_e), $X_1, \ldots, X_n$'lerin her biri F_e dağılımına sahip bağımsız rasgele değişken olmak üzere,

$$\text{OHK}(F_e) = B_{F_e}\left[\left(\sum_{i=1}^{n} \frac{X_i}{n} - \overline{x}\right)^2\right]$$

olur.

Örnek 8e Örnek 8a'daki dizgede, bir müşterinin uzun dönemde harcayacağı ortalama süreyi tahmin etmek isteyelim. Yani, i'inci sırada dizgeye giren müşterinin harcadığı süre W_i, $i \geqslant 1$ olmak üzere,

$$\theta \equiv \operatorname*{erey}_{n \to \infty} \frac{W_1 + W_2 + \cdots + W_n}{n}$$

değerini istiyoruz.

- **Uyarı Yenileyici** Yaklaşım Yukarıdaki çözümlemede, her günün bağımsız olarak aynı olasılık yasasına göre geçtiği varsayılmaktadır. Bazı uygulamalarda, aynı olasılık yasası dizgeyi sabit değil rasgele uzunluktaki döngülere göre günleri tanımlar.

Örneğin, bir müşterinin dizgede harcayacağı ortalama süre, θ, bir döngü içinde tüm gelenlerin dizgede harcadıkları sürelerin toplamı D ve bu gelişlerin sayısı N olmak üzere, $\theta = B[D]/B[N]$ olur. Eğer k tane döngü üretirsek, θ için tahmin edicimiz yine $\sum_{i=1}^{k} D_i / \sum_{i=1}^{k} N_i$ olarak kalır. Ayrıca, bu tahminin hata kare ortalaması, aynı yukarıda olduğu gibi özcül yöntem yaklaşımı kullanılarak yaklaşıklanabilir.