BMÜ-421 Benzetim ve Modelleme
STOKASTİK ÜRETEÇLER

İlhan AYDIN

RASGELE SAYI ÜRETEÇLERİ

- Deterministik terimler ile doğayı tanımlamak geleneksel bir yoldur.
- Doğa ve mühendislik sistemleri kesin olarak tahmin edilebilir bir tarzda değildirler.
- Sistemler genelde gürültü içerir bu yüzden bir sistemi gerçekçi modellemek için rastgeleliğin bir derecesi modele eklenmelidir.
- Olay tahmin edilmese bile sonraki olayların nasıl dağıtılacağı tahmin edilebilir.
- Verilen veriden ortalama, standart sapma ve benzeri hesaplamaların yapıldığı geleneksel istatistik analizinden farklı olarak bu bölümde ön tanımlı istatisiklere sahip veri kümesi üretme işleminden bahsedilecektir.

RASGELE SAYI ve DEĞİŞKEN ÜRETİMİ

- Gerçek sistemlerin olasılıklı stokastik davranışı her zaman düzgün (uniform) dağılımla açıklanamaz.
- Bir sistem içinde karşılaşılan stokastik işlemler uniform dağılımdan daha çok diğer teorik dağılımlarla (üstel, normal, gamma v.b.) açıklanabilmektedir.
- Bu nedenle uniform dağılımdan [0,1] aralığında elde edilen rassal sayıların teorik veya ampirik dağılımlara dönüştürülmesi gerekir.
- Bunun için bir dönüşüm tekniği kullanılarak 0-1 aralığında düzgün dağılımdan üretilen **rassal sayı** istenilen dağılım türünden bir **rassal değişkene** dönüştürülür.

RASTGELE SAYI:

- Herhangi bir dağılımdan rassal değişken üretmek veya bir rassal süreç üretmek için U(0,1) rassal değikenleri gereklidir. Bu nedenle kullanılan bilgisayarda istatistiksel olarak güvenilir bir rassal sayı üreteci olmalıdır. Eğer yoksa bir alt program olarak hazırlanıp yüklenebilir.
- Stokastik faaliyetleri konu alan benzetim modellerinde , olasılık dağılımlarından rassal değişken üretmek için rassal sayılar gereklidir. Bu nedenle bazı yazarlar MONTE-CARLO yöntemini , rassal sayılara dayalı deneylerle uğraşan deneysel matematiğin bir dalı olarak tanımlarlar.

RASSAL SAYI ÜRETEÇLERİNDEN İSTENİLEN ÖZELLİKLER:

- Rassallık
- Büyük Period
- Yeniden Üretilebilirlik (Reproducibility)
- Hesaplama Etkinliği

RASSAL SAYI ÜRETİMİ İÇİN TEKNİKLER

1) ORTA KARE YÖNTEMİ

- 1916'da Von Neumann ve Metropolis tarafından önerilen "ORTAKARE" yöntemidir
- Bu yöntemde , (m) basamaklı ve genellikle tek olan bir sayı başlangıç değeri olarak alınır
- İkinci aşamada, bu sayının karesi alınarak bulunan sayının ortasındaki m kadar basamaklı sayı alınır
- Bu bir rassal sayı olarak kayıt edilir
- Tekrar bu rassal sayının karesi alınır ve yine ortadaki m basamaklı sayı bir rassal sayı olarak kaydedilir
- Bu işlem , istenilen sayıda rassal sayı elde edilinceye kadar devam eder.

Örnek:

$X_0 = 5497$ olarak seçilsin.

$X_0^2 = (5497)^2 = 30.217.0,09 \Rightarrow$ **$X_1 = 2170$**

$U_1 = 0.2170$

$X_1^2 = (2170)^2 = 4.708.900 \Rightarrow$ **$X_2 = 7089$**

$U_2 = 0,7089$

$X_2^2 = (7089)^2 = 50.253.921 \Rightarrow$ **$X_3 = 2539$**

$U_3 = 0,2539$

Bu tekniğin dezavantajları ;

- İlk sayı ve dizinin tekrar uzunluğu arasındaki ilişkiyi (peryod) önceden bilmek mümkün değildir. Çoğu kez tekrar uzunluğu kısadır
- Elde edilen sayılar rassal olmayabilir
- Yani dizide dejenerasyon söz konusu olabilir.
- Bu yöntemle belirli bir sayı aritmetik işleme başlangıç değeri (seed) olarak verilmekte ve buna bağlı olarak bir sayı hesaplanmaktadır
- Hesaplanan sayı , bu kez başlangıç değeri olarak alınmakta ve yeni bir sayı üretilmektedir
- Böylece her üretilen sayıdan yeni bir sayı üretilerek bir sayı dizisi elde edilmektedir

TEKDÜZE DAĞITIMLI RASTGELE SAYILAR

- Dil derleyicileri [0,1] aralığında tekdüze dağılımlı rastgele sayılar için olanak sağlar.
- Böyle yordamlar U [0,1] üreteçleri olarak bilinir.
- Örneğin; BASIC dilinde RND çağrısı 0<=x<=1 aralığında bir x kesiri döndürecektir.
- Kesin konuşmak gerekirse, bu ayrık bir rastgele değişkendir.
- Fakat pratikte sürekli olduğu varsayılır
- 100 defa RND fonksiyonunu çağırırsanız kabaca %10'u 0 ile 0.1 arasında, %10'u 0.1 ile 0.2 arasında vb. dağılımlar oluşacaktır.

TEKDÜZE DAĞITIMLI RASTGELE SAYILAR

- Tek düze rastgele sayı üreteçlerinin çoğu LCG (LinearCongruentialGenerators) – Lineer Eşleşiksel Üreteçler – şeklindedir.
- Bunlar genelde deterministik olup bir algoritmaya dayalıdır.
- LCG, tahmin edilemez gibi görünen bir dizi sayılar oluşturur.
- Başlamak için bir ilk değer çekirdeğe Z_0 ihtiyaç duyar.
- Bu çekirdek ve Z_k dizisinin ardışıl terimleri bir LCG formülüne uygulanır.
- Ardından, Z_k, $0 \leq U_k \leq 1$ aralığında bir U_k çıkışına normalize edilir.
- Yani,

$$Z_0 = \text{"çekirdek"}\,, \quad Z_{k+1} = (aZ_k + c)mod(m)$$

$$U_k = \frac{Z_k}{m}$$

a : çarpan, c: artım ve m: genlik

TEKDÜZE DAĞITIMLI RASTGELE SAYILAR

- Örnek: a=5, c=3, m=16 ve Z_0= 7 değerleri ile LCG kullanarak oluşturulan sayı dizisini belirleyelim.

$$U_0 = \frac{Z_0}{m} = \frac{7}{16} \approx 0.437$$

$$Z_{k+1} = (5Z_k + 3)mod(16)$$

Z0=7➔ Z1=(5*7+3) mod 16=6 U1=6/16=0.375 olur.

- Benzer şekilde k=1 için Z2=1 ve U2=0.062 elde edilir.
- Burada Zk m ile bölünme sonucu elde edildiğinden, sadece m adet kalan vardır.
- Dolayısıyla bu örnekte maksimum 16 rastgele sayı mümkündür.
- Büyük m değerleri iyi bir seri elde etmek için gereklidir.
- m adet tekrar için m farklı sayının oluştuğu durumda seçilen LCG'nin tam periyoda sahip olduğu söylenir.
- Bu her bir Zk bir kez tekrar ettiği için tam periyot oluşmaktadır. Yukarıda verilen örnek tam periyoda sahip olup elde edilen rastgele sayılar aşağıdaki tabloda verilmiştir.

LCG ile oluşturulmuş sözde rastgele dizi		
k	Z_k	U_k
0	7	0.437
1	6	0.375
2	1	0.062
3	8	0.500
4	11	0.688
5	10	0.625
6	5	0.313
7	12	0.750
8	15	0.938
9	14	0.875
10	9	0.563
11	0	0.000
12	3	0.188
13	2	0.125
14	13	0.813
15	4	0.250

Dizinin ilk 16 elemanı tablodaki gibidir.

m tekrarlı bir durum için, m farklı rastgele sayı oluştuğunda LCG seçimi tam periyoda sahiptir.

Z_k nın bir tekrarında tam bir döngü izler.

Buradaki, LCG, tam periyoda sahiptir.

Hull-Dobell Teoremi

- Parametrelerin seçiminde Hull-Dobell teoremi oldukça kullanışlıdır.
- Bu teorem tam periyodu elde etmek için gerekli ve yeterli şartları sağlar.
- LCG ancak ve ancak aşağıdaki üç şartı sağlarsa tam periyoda sahiptir.
  I. a ve c asal olmalı
  II. m sayısının bölünebildiği bütün asal sayılara a-1 de bölünebilmelidir.
  III. Eğer m dörde bölünüyorsa a-1 de 4'e bölünebilir.
- Önceki örnekte
  - 5 ve 3 asal olduğu için şart (I),
  - m=16 olduğundan 16 sadece 2 asal sayısına bölünür ve a-1=5-1=4 de 2 ye bölünür(şart II).
  - 16 dörde bölünmekte ve a-1 de dörde bölünmektedir (şart III).
- Bütün şartlar sağlandığı için tam periyoda sahiptir.
- Bir bilgisayar uygulaması, bu algoritmayı donanım aşamasında ele alır. Çünkü, işlemler hesaplama ve hız odaklıdır.
- İşlem makineye shiftregister kullanılarak yaptırılır.
- m, 2'nin kuvveti şeklinde alınır.

TEKDÜZE DAĞITIMLI RASTGELE SAYILAR

Örnek: Önceki örnekteki problemi düşünelim. Değişkenler a=5, c=3, ve $m = 16 = 2^4$. Dolayısıyla LCG 4-bit shiftregister ile tam sayıları gösterebilir. R=[$r_{-1} \quad r_{-2} \quad r_{-3} \quad r_{-4}$].

Register içeriği 4 bit olacaktır.

$Z_6 = 5$ olduğundan R:[0101] dir

Z_7'yi elde etmek için $5Z_6 + 3 = \text{R}: [1 \;\; 1100] = 28$

Burada baştaki 1 shift-register 4 bit olduğundan kaybedilir.

$$28 \bmod(16) = 12 = \text{R}: [1100]$$

R←5R+3: [1 1 1 1] elde edilir Z8=15 olur.

İkili nokta uygulandığında $(0.1100)_2 = 0.75$

- Gerçek bilgisayarlarda farklı ölçüde üreteçler vardır.
- IBM'in RANDU üreteçleri, $a = 2^{16} + 3$, $c = 0$ ve $m = 2^{31}$ sahiptir.

U[0,1] ÜRETEÇLERİN İSTATİSTİKSEL ÖZELLİKLERİ

Üreteçlerin İstatistiksel Özellikleri

- Donanım hesaplanabilirliği için seçilen mod işlemi ve geniş bir peryoda sahip olmanın yanı sıra bir U[0,1] üreteci istatistiksel anlamda iyi davranmalıdır.
- Şu iki özelliğin sağlanması önemlidir:
  - Üreteç tekdüze olmalı: Herhangi bir L uzunluk aralığında oluşan sayıların miktarı, diğer bir L uzunluk aralığında oluşan miktara yakın olmalı.
  - Dizi bağımsız olmalı: Özellikle, herhangi bir sayı bir sonrakine etkisini göstermemelidir. Aksi halde dizi boşluk veya gruplama eğilimi gösterir.
- Üreteçleri test etmek için teorik ve deneysel araçlar vardır.
- Birinci özelliği test etmek için chi-square (Ki-Kare) testi uygulanır.
- Ki-Kare testi; beklenen frekans değerler ile gözlenen frekans değerlerinin karşılaştırılıp, aradaki uyuma bakılmasıdır.

Frekans Dağıtım Tablosu			
Aralık sayısı k	Aralık	Deneysel frekans f_k	Beklenilen frekans e_k
1	$[0, 1/m]$	f_1	e_1
2	$[1/m, 2/m]$	f_2	e_2
3	$[2/m, 3/m]$	f_3	e_3
.	.	.	.
.	.	.	.
.	.	.	.
m	$[(m-1)/m, 1]$	f_m	e_m

- Bu test için, FDT (Frequency Distribution Table) – Frekans Dağıtım Tablosu- faydalanılır.

- m rastgele sayı oluşturularak ve her birini bir m sınıfına atayarak $f_1, f_2, \ldots f_m$ frekansları çizelgeye geçirilir.

- Her bir sınıf için beklenen $e_k = \frac{n}{m}$ frekansı ile karşılaştırılır.

$$\chi^2 = \sum_{k=1}^{m} \frac{(f_k - e_k)^2}{e_k}$$

$$= \frac{m}{n} \sum_{k=1}^{m} (f_k - \frac{n}{m})^2$$

v=m-1 bağımsızlık derecesidir.

Üreteçlerin İstatistiksel Özellikleri

- Örnek: SNAFU olarak isimlendirilen U[0,1] üreteci 100 sayı üretilerek test edilmiş ve frekansları sayılmıştır. Frekans değerleri aşağıda verilmiştir.

  $0.00 \le x < 0.25$
  $0.25 \le x < 0.50$
  $0.50 \le x < 0.75$
  $0.75 \le x < 1.00$

- Sonuçlar f1=21, f2=31, f3026, f4=22 şeklindedir. Üretecin uniform olup olmadığını bulunuz?

n=100 m=4 sınıf var. n/m=25 sayı her sınıfta olmalıdır. Ki-kare testi ile aşağıdaki gibi bir sonuç elde edilir.

$$\chi^2 = \frac{4}{100}[(21-25)^2 + (31-25)^2 + (26-25)^2 + (22-25)^2] = 2.48.$$

Bağımsızlık derecesi v=4-1=3 χ^2 *değeri* $\alpha = 95\% \chi_c^2$ =7.81(Appendix F) olduğu ki-kare tablosundan bulunabilir.

$\chi^2 < \chi_c^2$ olduğundan uniform olduğu söylenebilir.

Appendix F THE CHI-SQUARE DISTRIBUTION FUNCTION

Values of x for given $1 - F(x)$ with ν degrees of freedom

	Complemented distribution, $1 - F(x)$			
Degrees of freedom ν	0.10	0.05	0.02	0.01
1	2.706	3.841	5.412	6.635
2	4.605	5.991	7.824	9.210
3	6.251	7.815	9.837	11.341
4	7.779	9.488	11.688	13.277
5	9.236	11.070	13.388	15.086
6	10.645	12.592	15.033	16.812
7	12.017	14.067	16.622	18.475
8	13.362	15.507	18.168	20.090
9	14.684	16.919	19.679	21.666
10	15.987	18.307	21.161	23.209
11	17.275	19.675	22.618	24.725
12	18.549	21.026	24.054	26.217
13	19.812	22.362	25.472	27.688
14	21.064	23.685	26.873	29.141
15	22.307	24.996	28.259	30.578
16	23.542	26.296	29.633	32.000
17	24.769	27.587	30.995	33.409
18	25.989	28.869	32.346	34.805
19	27.204	30.144	33.687	36.191
20	28.412	31.410	35.020	37.566
21	29.615	32.671	36.343	38.932
22	30.813	22.924	37.659	40.289
23	32.007	35.172	38.968	45.638
24	33.196	36.415	40.270	42.980
25	34.382	37.652	41.566	44.314
26	35.563	38.885	42.856	45.642
27	36.741	40.113	44.140	46.963
28	37.916	41.337	45.419	48.278
29	39.087	42.557	46.693	49.588
30	40.256	43.773	47.962	50.892

GENERATION OF NON-UNIFORM RANDOM VARIATES

TEKDÜZE OLMAYAN RASGELE DEĞİŞKENLERİN ÜRETİMİ

TEKDÜZE OLMAYAN RASGELE DEĞİŞKENLERİN ÜRETİMİ

- İstatistiksel dağıtımda, isteğe bağlı sayıları oluşturabilmek önemlidir. Bunu yapabilmek için bazı bilinen algoritmalar vardır.
  - Formül Metodu
  - Ret Metodu
  - Konvolüsyon Metodu

Formül Metodu

- Örneğin, belirli rastgele rassal değişkenler U[0,1] üretecini içeren bir formül olarak basit bir şekilde gösterilebilir.
- RND bir üreteç olarak kabul edilirse, [a, b] isteğe bağlı aralığında tekdüze rastgele X sayılarını bulmak mümkündür.

$$X = (b - a)RND + a$$

- U'nun minimum değeri 0 olduğundan X'in minimum değeri;

$$X = (b - a)(0) + a = a$$

- U'nun maksimum değeri 1 olduğundan X'in maksimum değeri;

$$X = (b - a)(1) + a = b$$

Formül Metodu

- Formüller tek-düze olmayan dağılımlar için de vardır.
- Normalize bir Gaussian rastgele değişkenlerin oluşumu şu formül ile sağlanır:

$$Z = \sqrt{-2\ln(RND)}\cos(2\pi RND).$$

- Z ortalaması 0, standart sapması 1 olan rastgele bir değişkendir.
- Doğrulaması Ki-Kare testi ile yapılabilir.
- Rasgele bir değişken için elde edilen formüllerden biri ters çevrim (inversion) metodudur.
- Bu metot F dağılım fonksiyonunun tersinin ve hesaplanmasını gerektirir.

$$X = F^{-1}(RND)$$

- F^{-1}, F dağıtıklaştırma fonksiyonunun tersidir.

Örnek

- μ ortalamaya sahip bir exponansiyel dağıtık rasgele değişenleri oluşturan formül elde edelim.
- Bir exponansiyel rasgele değişken için yoğunluk fonksiyonu;

$$f(x) = \begin{cases} \frac{1}{\mu} e^{-x/\mu}, & x \geq 0, \\ 0, & x < 0 \end{cases}$$

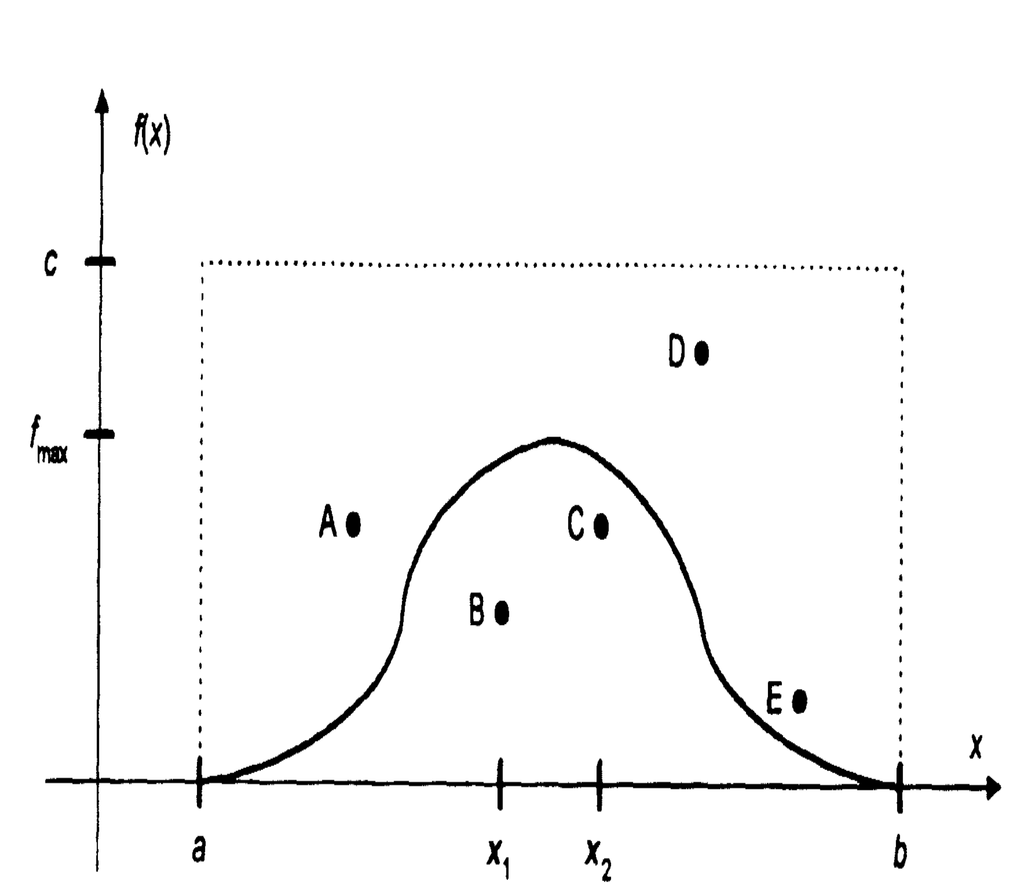
- $$F(x) = \int_{-\infty}^{x} f(t)dt = \int_{0}^{x} \frac{1}{\mu} e^{-t/\mu} dt$$

$$= 1 - e^{-x/\mu}, \quad x \geq 0$$

$$F^{-1}(x) = X = -\mu \ln(RND)$$

Rejection (Red) Metodu

- Dağıtık fonksiyonun her zaman tersini bulabilmek mümkün değildir.
- Bu gibi durumlar için Rejection (Red) Metodu kullanılır.



$$x = a + (b - a)RND$$
$$y = c * RND$$

Red Metodu Algoritma

- [1] $x = a + (b - a) * RND$

  $y = c * RND$

  $if\ y > f(x)\ \ then\ goto\ [1]$

  $print\ x$

- A,D,E noktaları, hedefin içinde olmasına rağmen eğrinin dışında olduğundan önemsenmez.
- B ve C ‘nin $x_1\ ve\ x_2$ noktaları iki rastgele değişkendir.

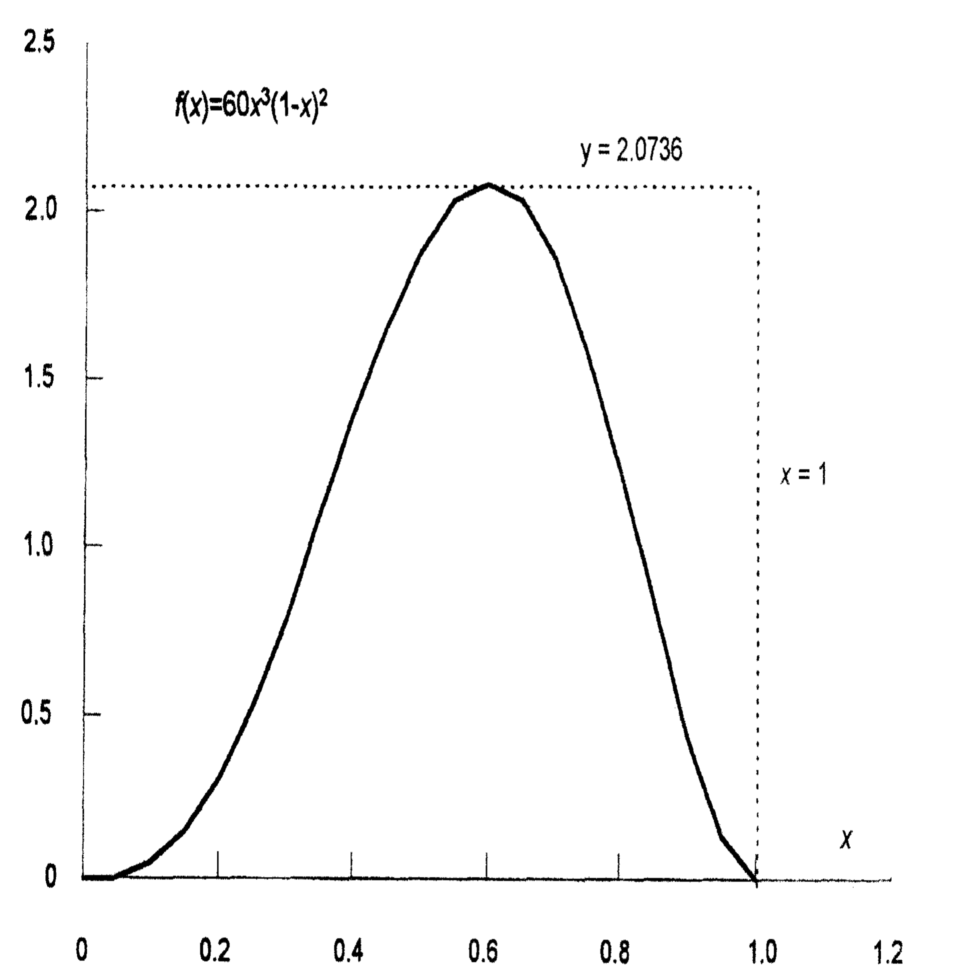
Örnek

- $a = 3, b = 2$ ile Beta dağıtıklaştırma için 100 rastgele değişken oluşturan red metodunu uygulayalım.

- Gerekli Beta fonksiyonu;

$$f(x) = \frac{\Gamma(7)}{\Gamma(4)\Gamma(3)} x^3 (1-x)^2 = 60x^3(1-x)^2,$$

$$0 \leq x \leq 1$$



- Optimum y aralığını bulabilmek için maksimum $f(x)$ bulunmalıdır.

- Maksimum $f(x)$ için;

$$f'(x) = 0$$

$$f'(x) = -120x^3(1-x) + 180x^2(1-x)^2$$

$$= 60x^2(1-x)(3-5x) = 0$$

Çözüm Kümesi: $x = 0, x = 1, x = 0.6$

- $x = 0\ ve\ x = 1$ uç noktalardır.

- $f(0.6) = 60 * (0.6)^3 * (1 - 0.6)^2$
  $= 2.0736$

- Hedefin üst kenarı 2.0736'dır.

- Algoritma;

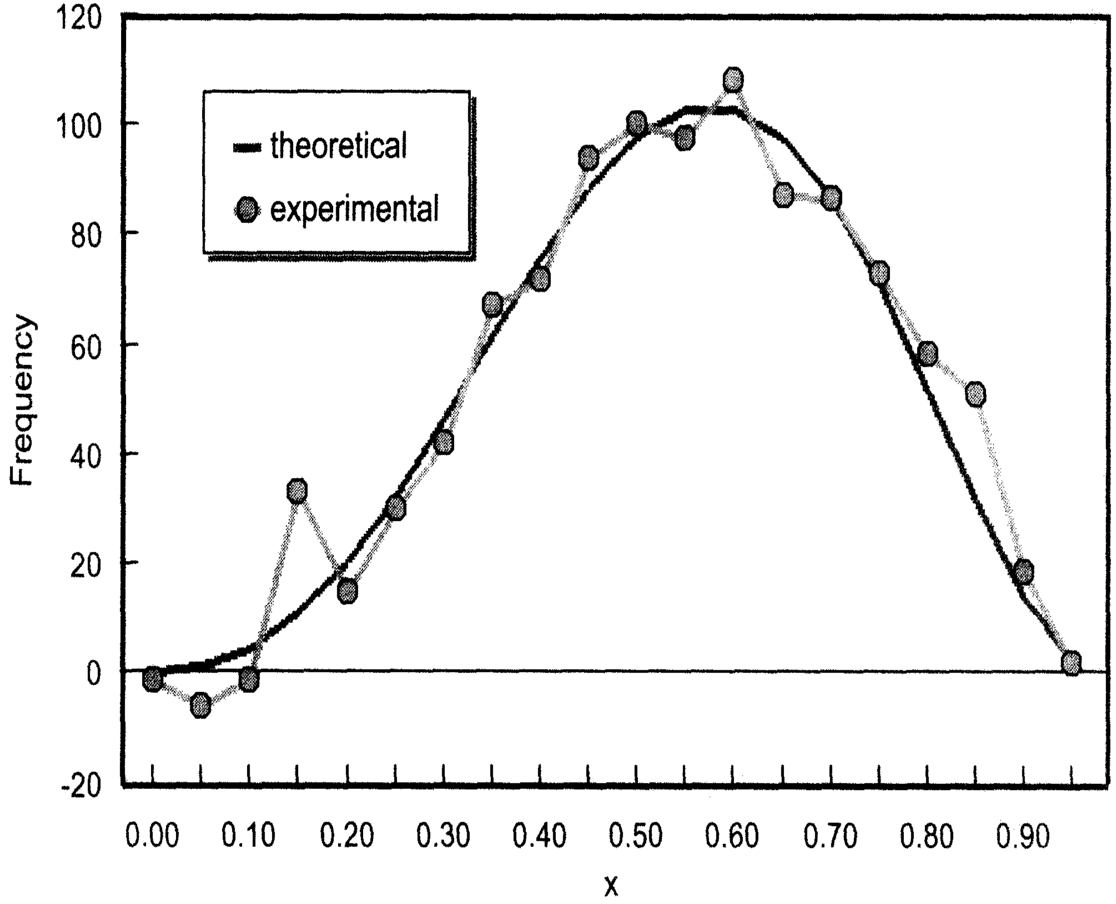
$$
\begin{gathered}
for\ \ i = 1\ \ to\ 100 \\
[1]x = RND \\
y = 2.1\ * RND \\
if\ \ y > 60x^3\ (1 - x)^2\ then\ goto\ [1] \\
print\ x \\
next\ i
\end{gathered}
$$

- Algoritmanın geçerliliği Frekans Dağıtım Tablosu ile sınanabilir.

- $\Delta x = 0.05$ ve $n = 1000$ için değerler şu şekildedir.

Frekans Dağıtım Tablosu

Frekans Dağıtım Tablosu		
Aralık	Deneysel frekans	Beklenilen frekans
[0.00, 0.05]	4	0.09
[0.05, 0.10]	5	1.15
[0.10, 0.15]	2	4.62
[0.15, 0.20]	5	11.07
[0.20 0.25]	23	20.64
[0.25 0.30]	38	32.87
[0.30 0.35]	46	46.95
[0.35 0.40]	64	61.78
[0.40 0.45]	69	76.06
[0.45 0.50]	86	88.49
[0.50 0.55]	110	97.77
[0.55 0.60]	111	102.81
[0.60 0.65]	94	102.77
[0.65 0.70]	81	97.22
[0.70 0.75]	78	86.26
[0.75 0.80]	85	70.55
[0.80 0.85]	44	51.54
[0.85 0.90]	52	31.49
[0.90 0.95]	5	13.62
[0.95 1.00]	6	2.23



theoretical
experimental
Frequency
120
100
80
60
40
20
0
-20
0.00
0.10
0.20
0.30
0.40
0.50
0.60
0.70
0.80
0.90
x

- Hedef bölgesinin dikdörtgen olmasına ihtiyaç yoktur. Amaç yoğunluk fonksiyonunu çevrelemektir.
- İterasyon sayısı arttıkça uygulama yavaşlayacaktır.
- Yoğunluk fonksiyonu karmaşıklaştıkça çevrelemek zorlaşır. Bu yüzden uç noktalar ihmal edilebilir.

Convolution (Konvolüsyon) Metodu

- Bağımsız ve özdeş dağıtılan (independentandidenticallydisributed - IID) $X_1, X_2, \ldots\ X_n$ rasgele değişkenlerinin toplamı olan X değişkenidir.
- Eğer $X_i,\ i = 1, 2, \ldots, n$ için aynı yoğunluk fonksiyonu $f_i(x)$'e sahip ise X'in yoğunluk fonksiyonu $f(x), n$ tabanlı yoğunluk fonksiyonlarının her biri için konvolüsyondur.

Konvolüsyon Metodu

- Yani;

$$X = \sum_{k=1}^{n} X_k \quad ise, f(x) = f_1(x) \otimes f_2(x) \otimes \ldots \otimes f_n(x)$$

$f_i(x), X'$in yoğunluk fonksiyonu

$\otimes$, konvolüsyon ifadesidir.

Konvolüsyon Metodu

$$f_1(x)\otimes f_2(x) = \int_{-\infty}^{\infty} f_1(\lambda) f_2(x-\lambda)\,d\lambda$$

- Rasgele değişken kendini, $X = \sum_{k=1}^{n} X_k$ n tane IID değişkenine ekleyerek bulur.

- Konvolüsyon metodu için özel bir durum, m $-$Erlang dağıtımıdır.

m –Erlang Dağıtımı

- m adet IID exponansiyel rasgele değişkenin toplamı olarak tanımlanır.
- Bu dağıtımın ortalaması;

$$\mu = E\left[\sum_{k=1}^{m} X_k\right] = \sum_{k=1}^{m} E[X_k] = \frac{m}{\lambda}.$$

λ , exponansiyel dağıtımın ortalamasının matematiksel karşıtıdır.

m –Erlang Dağıtımı

- Rasgele bir m –Erlangdeğişkeni oluşturma algoritması;

$$x = 0$$
$$for\ k = 1\ \ to\ m$$
$$x = x - \mu \ln(RND)/m$$
$$next\ k$$
$$print\ x$$

Örnek

- Ortalaması 5 olan 1000 elemanlı $2-$Erlang dizisi oluşturalım ve Ki-Kare testi ile kıyaslama yapalım.

- Çözüm:

$2-$Erlang dağıtımı $\alpha = 2$ ile Gamma dağıtımının özel bir durumudur.

- Ortalama 5 ise,

$\frac{2}{\lambda} = 5, \qquad \lambda = 0.4$ olur.

$2-$Erlang dağıtımı için yoğunluk fonksiyonu;

$$f(x) = \frac{4}{25} x e^{-2x/5} \quad , \quad x \geq 0 \text{ olur.}$$

- Konvolüsyon metoduna göre, bir Erlang rasgele değişkeni $-2.5\ln(RND)$ ve $-2.5\ln(RND)$'nin toplamıdır.
- Cebirsel karşılığı $-2.5\ln(RND * RND)$

- Sonuçların doğrulanması için gerekli $n = 1000$ rasgele değişken frekans dağıtım tablosunda özetlemiştir.

- Her bir aralıktaki beklenilen frekans;

$$E_{[a,b]} = n\int_a^b f(x)dx = \frac{4n}{25}\int_a^b xe^{-2x/5}dx$$

$$= n[e^{-2x/5}(1-\frac{2x}{5})]_a^b$$

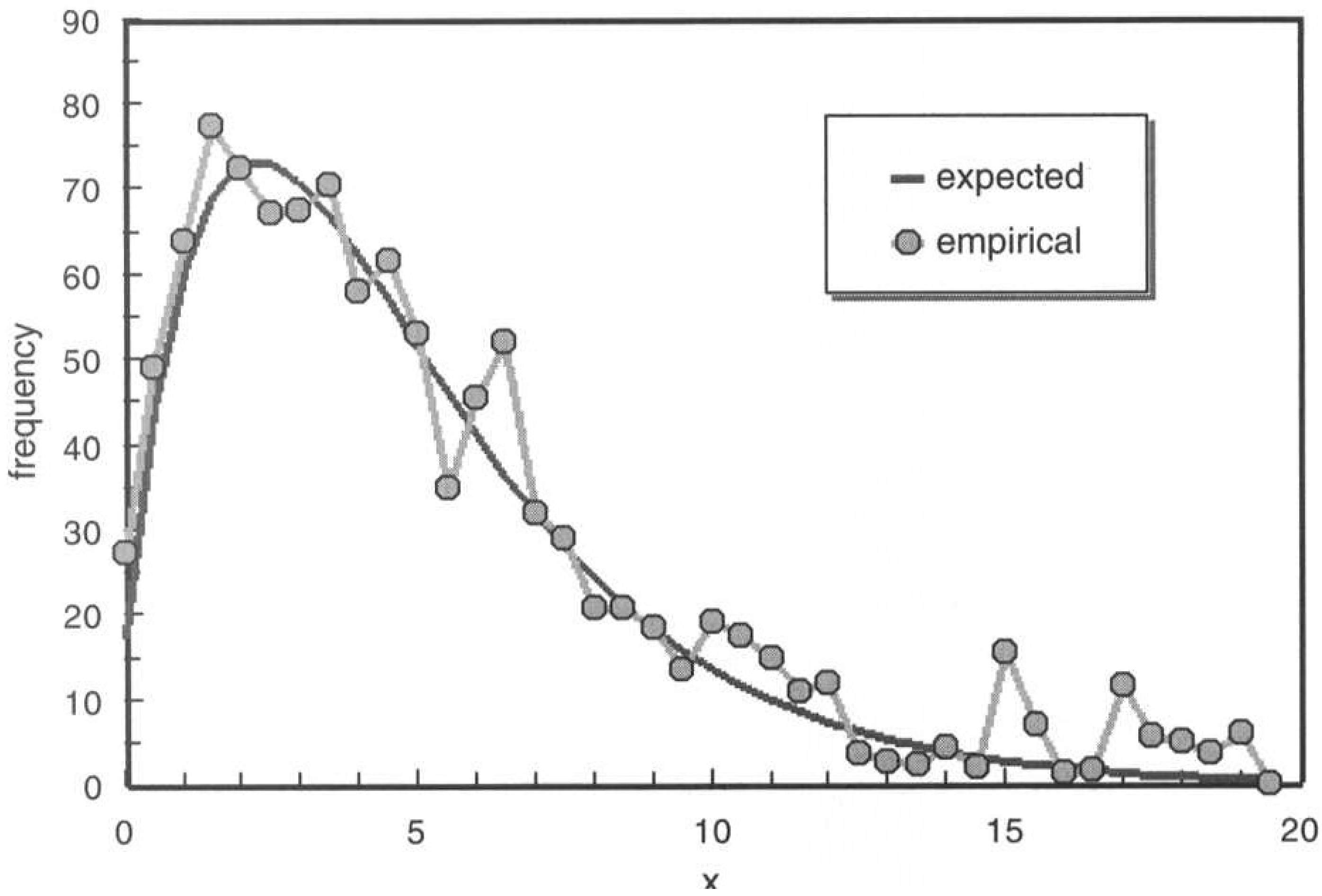
$$= n\left[\left(e^{-2b/5} - e^{-2a/5}\right) + \frac{2}{5}(ae^{-2a/5} - be^{-2b/5})\right].$$

Frekans Dağıtım Tablosu

Frekans Dağıtım Tablosu

Aralık	Deneysel frekans	Beklenilen frekans
[0.0, 0.5]	8	17.52
[0.5, 1.0]	37	44.03
[1.0, 1.5]	56	60.35
[1.5, 2.0]	64	69.31
[2.0, 2.5]	76	73.03
[2.5, 3.0]	64	73.13
[3.0, 3.5]	77	70.79
[3.5, 4.0]	78	66.90
[4.0, 4.5]	64	62.09
[4.5, 5.0]	49	56.83
[5.0, 5.5]	53	51.44
[5.5, 6.0]	46	46.13
[6.0, 6.5]	50	41.06
[6.5, 7.0]	35	36.31
[7.0, 7.5]	27	31.93
[7.5, 8.0]	29	27.95
[8.0, 8.5]	21	24.36
[8.5, 9.0]	22	21.15
[9.0, 9.5]	9	18.31
[9.5, 10.0]	25	15.80
[10.0, 10.5]	21	13.60
[10.5, 11.0]	9	11.68
[11.0, 11.5]	4	10.01
[11.5, 12.0]	6	8.56
[12.0, 12.5]	3	7.30
[12.5, 13.0]	5	6.22
[13.0, 13.5]	10	5.30
[13.5, 14.0]	10	4.50
[14.0, 14.5]	5	3.82
[14.5, 15.0]	5	3.24

Grafiksel Gösterim



GENERATION OF ARBITRARY RANDOM VARIATES

KEYFİ RASGELE DEĞİŞKENLERİN ÜRETİMİ

KEYFİ RASGELE DEĞİŞKENLERİN ÜRETİMİ

- Net bir yoğunluk fonksiyonunun bulunamadığı durumlar için kullanılır.
- Bunun yerine deneysel değişkenler kümesi kullanılır.
- Sistem için, tarihsel kayıtlar tespit edilir ve bu değerle ile aynı istatistiklere sahip rasgele değişkenler oluşturulur.

KEYFİ RASGELE DEĞİŞKENLERİN ÜRETİMİ

- Süreç iki aşamalıdır.
- Artan şekilde sıralanmış $\{x_1, x_2, x_3, \dots, x_n\}$ veri kümesini ele alalım.
- Birinci adım, parçalı-lineer ve sürekli yoğunluk fonksiyonunu elde etmektir.

$$F(x) = \begin{cases} 0, & x < x_1, \\ \dfrac{i-1}{n-1} + \dfrac{x - x_i}{(n-1)(x_{i+1} - x_i)}, & x_i \le x < x_{i+1}, i = 1, \dots n - \\ 1, & x \ge x_n \end{cases}$$

KEYFİ RASGELE DEĞİŞKENLERİN ÜRETİMİ

- İkinci adım, $F^{-1}(x)$'i bulmaktır.

$F(x)$ lineer olduğundan;

$$X = F^{-1}(x)$$

Örnek

- Bilinmeyen bir süreçten alınan sıralı rasgele değişkenler kümesi:

$\{1,2,4,5,7,7,9\}$ olsun.

Dağıtım fonksiyonunu ve tersini inceleyelim.

Çözüm

- Problem 7 adet veri noktasına sahiptir.

$x < 1$ için $F(x) = 0,$

$1 \leq x < 2$ için $i = 1,$

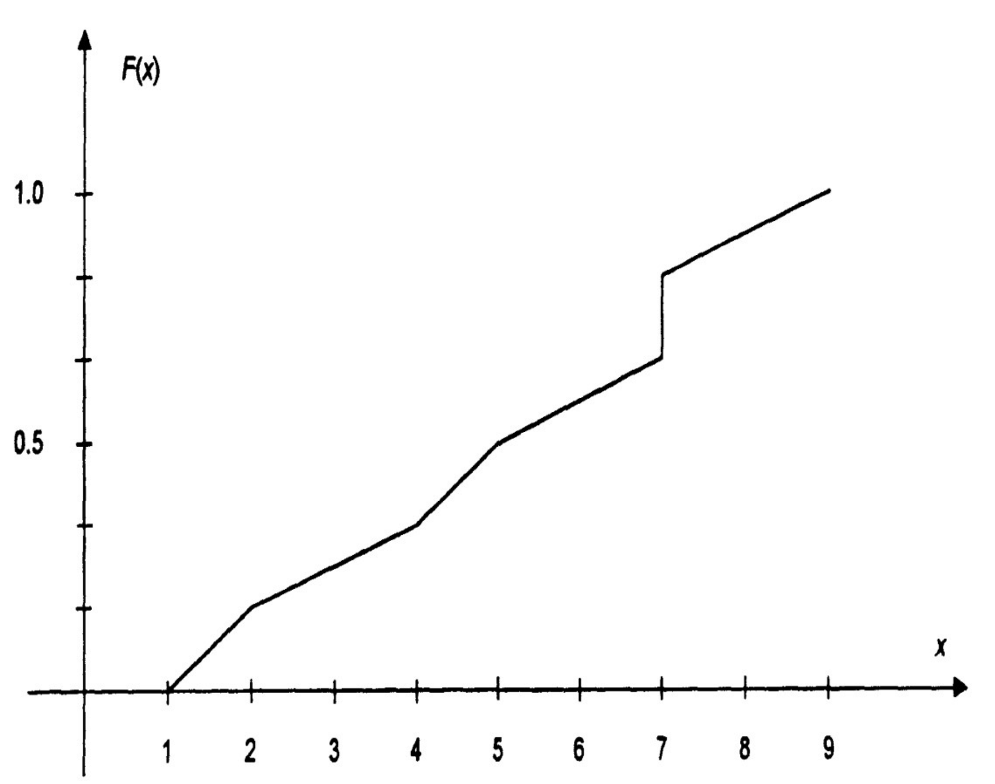
$$f(x) = \frac{1-1}{7-1} + \frac{(x-1)}{(7-1)(2-1)} = \frac{1}{6}(x-1)$$

$2 \leq x < 4$ için $i = 2,$

$$f(x) = \frac{2-1}{7-1} + \frac{(x-2)}{(7-1)(4-2)} = \frac{1}{12}x$$

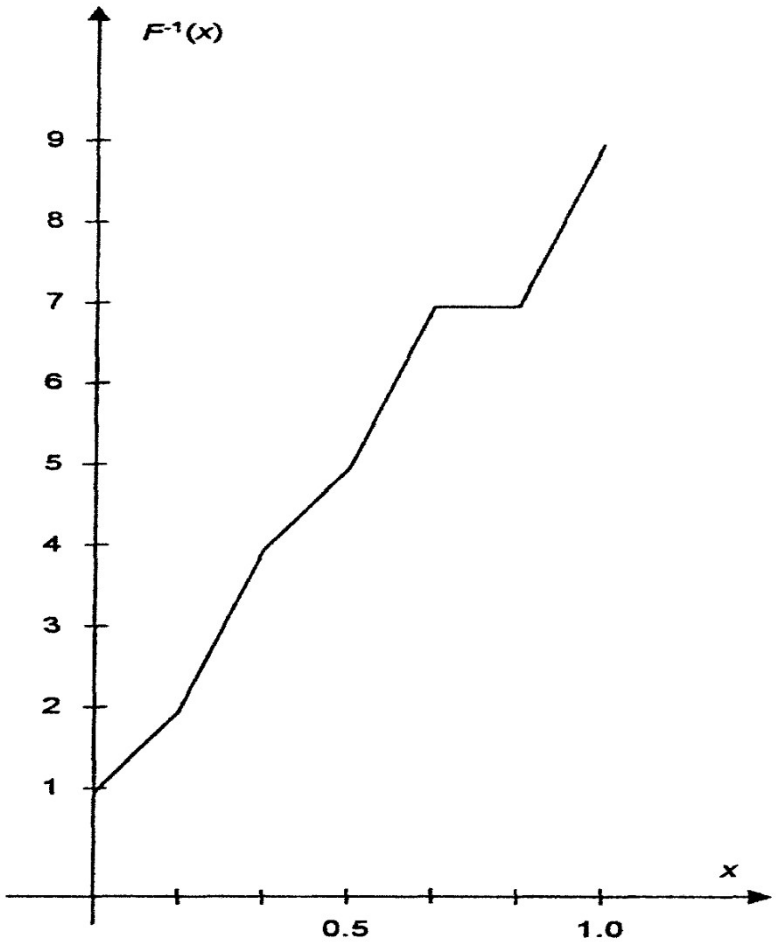
- Benzer şekilde diğer değerler için,

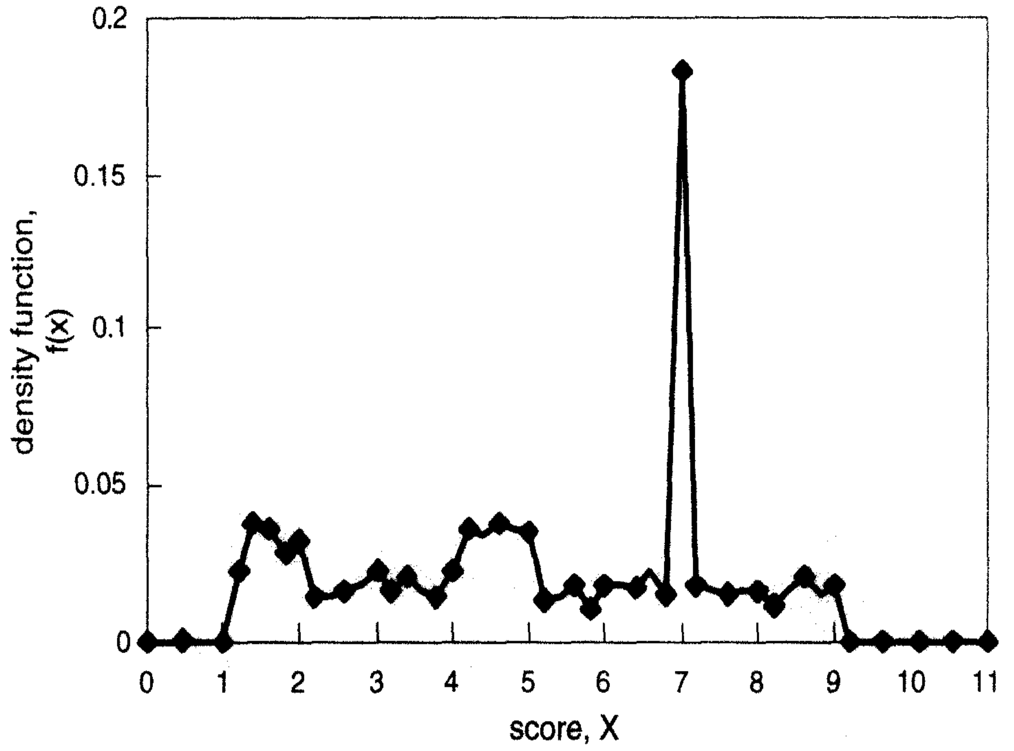
$$F(x) = \begin{cases} 0, & x < 1, \\ \frac{1}{6}(x-1), & 1 \leq x < 2, \\ \frac{1}{12}x, & 2 \leq x < 4, \\ \frac{1}{6}(x-2), & 4 \leq x < 5 \\ \frac{1}{12}(x+1), & 5 \leq x < 7 \\ \frac{1}{12}(x+3), & 7 \leq x < 9 \\ 1, & x \geq 9 \end{cases}$$


F(x)
1.0
0.5
x
1
2
3
4
5
6
7
8
9

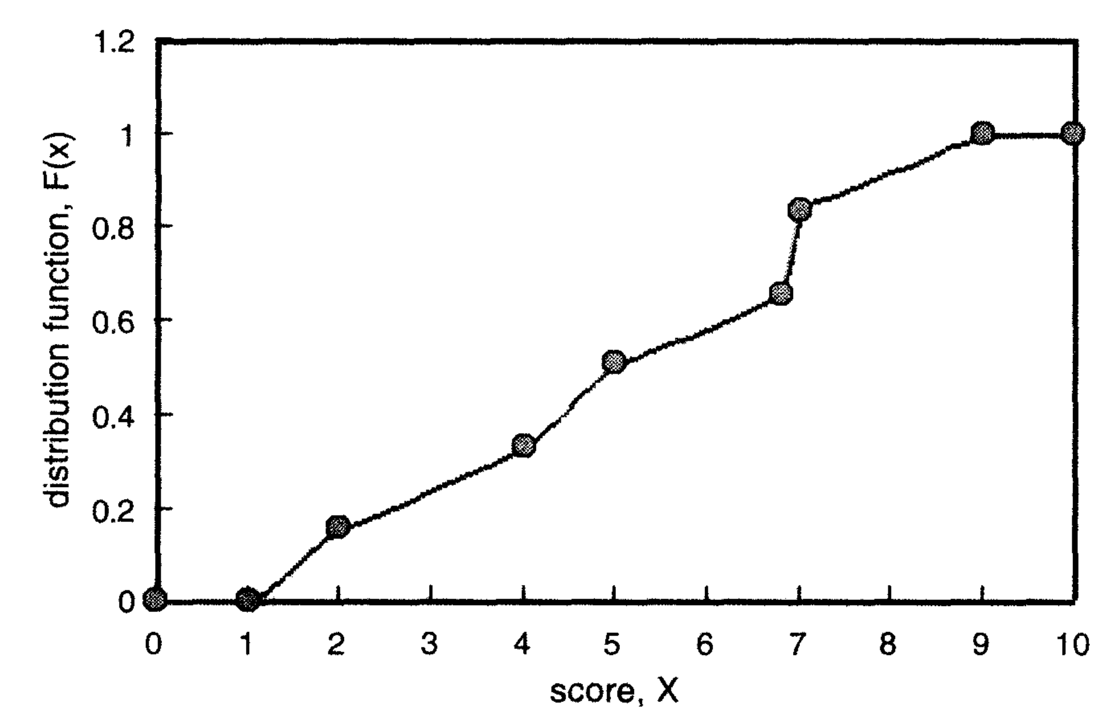
F^{-1} fonksiyonları

$$F^{-1}(x) = \begin{cases} tanımsız, & x < 1, \\ 6x + 1, & 0 \leq x < \frac{1}{6}, \\ 12x, & \frac{1}{6} \leq x < \frac{1}{3}, \\ 6x + 2, & \frac{1}{3} \leq x < \frac{1}{2}, \\ 12x + 1, & \frac{1}{2} \leq x < \frac{2}{3} \\ 7, & \frac{2}{3} \leq x < \frac{5}{6} \\ 12x - 3\,, & \frac{5}{6} \leq x < 9 \\ tanımsız, & x > 1 \end{cases}$$


$F^{-1}(x)$
9
8
7
6
5
4
3
2
1
x
0.5
1.0



0.2
0.15
0.1
0.05
0
density function, f(x)
0
1
2
3
4
5
6
7
8
9
10
11
score, X



1.2
1
0.8
0.6
0.4
0.2
0
distribution function, F(x)
0
1
2
3
4
5
6
7
8
9
10
score, X

RANDOM PROCESSES

RASTGELE SÜREÇLER

RASTGELE SÜREÇLER

- Dinamik bir sistemin simülasyonunu yaparken, sadece sistemin kendisini modellemek önemli değildir. Daha önemlisi giriş sinyalini elde etmektir.
- Çünkü çoğunlukla sistemin giriş sinyali belli değildir.
- Önceden kestirilemeyen bir sürecin etrafında yapısal bir dalgalanmaya sahiptir.

RASTGELE SÜREÇLER

- İyi modellerin çoğunda sistemin o anki değeri bir sonraki değeri muhteşem bir şekilde etkiler.
- Sinyalin olasılıksal olması tamamen rasgele olduğu anlamına gelmez.
- Süreçler, sinyaller koleksiyonudur.
- Bir sürekli zaman rasgele süreci $X(t)$ ile gösterilir.

RASTGELE SÜREÇLER

- Koleksiyon içindeki herhangi bir sinyal $x(t)$ ile gösterilir.
- Ortalama ve otokorelasyon rasgele süreçler için önemlidir.

$$\mu_x(t) = E[X(t)]$$

$$R_{xx}(t, \tau) = E[X(t)X(t+\tau)]$$

$$t = zaman, \tau = gecikme\ zamanı$$

RASTGELE SÜREÇLER

- Otokorelasyon, sinyal ile kendisinin τ zaman sonraki çarpımının beklenen değeri olarak görünür.
- $\mu_{X(t)} = E[X(t)]$ eğer bir sabit ise $wide - sense\ stationary$ (geniş algı sabiti) olarak isimlendirilir.
- Sabit özelliği önemlidir çünkü ortalama sürekli durum davranışı gösterir.

Örnek:

- [0,1] aralığında dört benzer sinyal içeren bir dizi ele alalım.

$$X(t) = \{2t + 1, t + 2, 3t + 2, 4t + 1\}.$$

Ortalamasını, ikinci momentini, varyansını ve otokorelasyonunu bulalım.

Ortalama

$$\mu = \frac{1}{4}[(4t + 1) + \ (2t + 1) + \ (t + 2) + \ (3t + 2)]$$

$$= \frac{1}{2}(5t + 3)$$

Moment

$$E[X^2(t)]$$

$$= \frac{1}{4}[(4t+1)^2 + (2t+1)^2 + (t+2)^2 + (3t+2)^2]$$

$$= \frac{1}{2}(15t^2 + 14t + 5)$$

Varyans

$$\sigma^2 = E[X^2(t)] - \mu^2$$

$$= \frac{1}{2}(15t^2 + 14t + 5) - \frac{1}{4}(5t + 3)^2$$

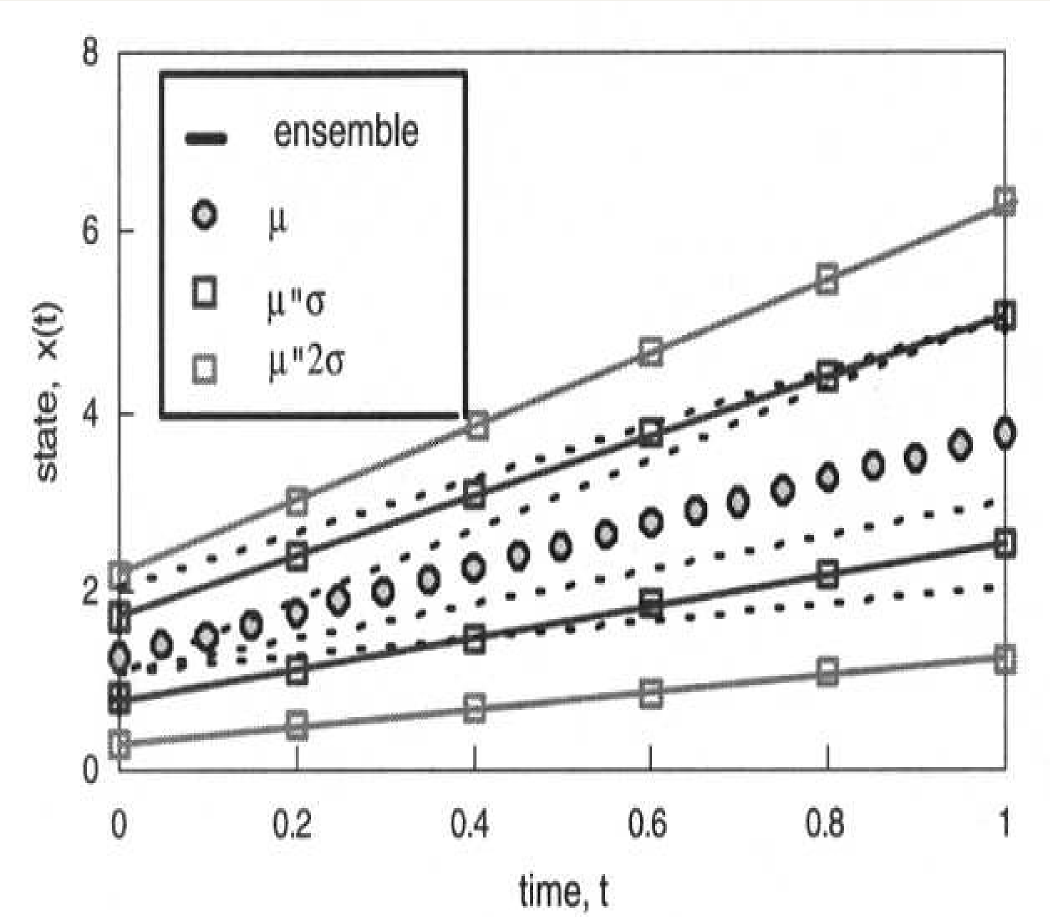
$$= \frac{1}{4}(5t^2 - 2t + 1)$$

Otokorelasyon

$$R_{xx}(t,\tau) = E[X(t)X(t,\tau)]$$

$$= \frac{1}{4}[(4t+1)(4t+4\tau+1) + (2t+1)(2t+2\tau+1)$$
$$+(t+2)(t+\tau+2) + (3t+2)(3t+3\tau+2)]$$

$$= \frac{1}{2}(15t^2 + 14t + 5 + 15\,t\tau + 7\tau)$$



ensemble
μ
μ"σ
μ"2σ
state, x(t)
time, t
0
2
4
6
8
0
0.2
0.4
0.6
0.8
1

- Realistik durumlarda her zaman, bir örnek kümesine sahip olunmaz.
- Bunun yerine süreci karakterize eden formüller veya rasgele değişkenler kaçınılmazdır.
- Karakterize işlemi, genellikle $\mu(t)$ ortalamasının ve $R_{xx}(t, \tau)$otokorelasyonundan oluşur.

CHARACTERIZING RANDOM PROCESSES

RASTGELE SÜREÇ KARAKTERİZASYONU

RASTGELE SÜREÇ KARAKTERİZASYONU

- Rasgele süreçler, $\{x_i(t)\}$ fonksiyonlarının veya sinyallerinin koleksiyonudur.
- Sinyal üretimi için açık bir formül yoktur.
- Sinyalin karakteristiklerinin üretebilmek önemlidir.
- Bu yüzden iki ana prensip vardır:

RASTGELE SÜREÇ KARAKTERİZASYONU

- Rastgele Değişken: Yoğunluk fonksiyonu ile sayıların bir dizisini oluşturmak

- Rastgele Süreç: Otokorelasyon ile sinyallerin bir dizisini oluşturmak

RASTGELE SÜREÇ KARAKTERİZASYONU

- Bir rastgele sürece ait özellikler;

  Ortalama

  Standart sapma

  Eşitlik

  Varyans

GENERATING RANDOM PROCESSES

RASTGELE SÜREÇLERİN OLUŞUMU

RASTGELE SÜREÇLERİN OLUŞUMU

- Bir yaklaşıma göre rastgele süreçler rastgele sinyallerdir.
- İki tip temel rastgele sinyal tipi vardır:
  - Düzenli
  - Aralıklı

RASTGELE SÜREÇLERİN OLUŞUMU

- Düzenli durumda;
  - Zaman, senkrondur. Her bir zaman vuruşunda bir sinyal üretilir.
  - Eşit aralıklı zaman artışları değer dizisini oluşturur.
- Aralıklı durumda;
  - Olaylar oluştuğunda sistem değeri değişmez.

RASTGELE SÜREÇLERİN OLUŞUMU

- Monolitik ve otonom sistemlerde, kararlar lokal olduğundan rastgele olmayan veya düzenli rastgele prosesler ile gösterilir.
- Bu süreçler sadece zamana bağlı olduğundan bu sistemlere zaman yürütümlü olarak bakarız.

RASTGELE SÜREÇLERİN OLUŞUMU

- Eğer kararlar sistem dahili olması yerine harici bir kaynaktan gelen olaylara göre alınıyorsa bu sistemler olay yürütümlü sistemlerdir.

- Veya sistem hem zaman hem de olay yürütümlü olabilir.

- Diğer bir bakış açısıyla sistemler iki türlüdür:

- Açık döngü sistemler; geri besleme kullanmazlar. Tüm kararlar otonomdur ve zamanla değişir.
- Kapalı döngü sistemler, harici sistem üyelerinden veya sensörlerden gelen gelen kararlardan etkilenir ve kesmeler kullanır.

- Açık döngü sistemler, zaman yürütümlü
- Kapalı döngü sistemler, olay yürütümlü

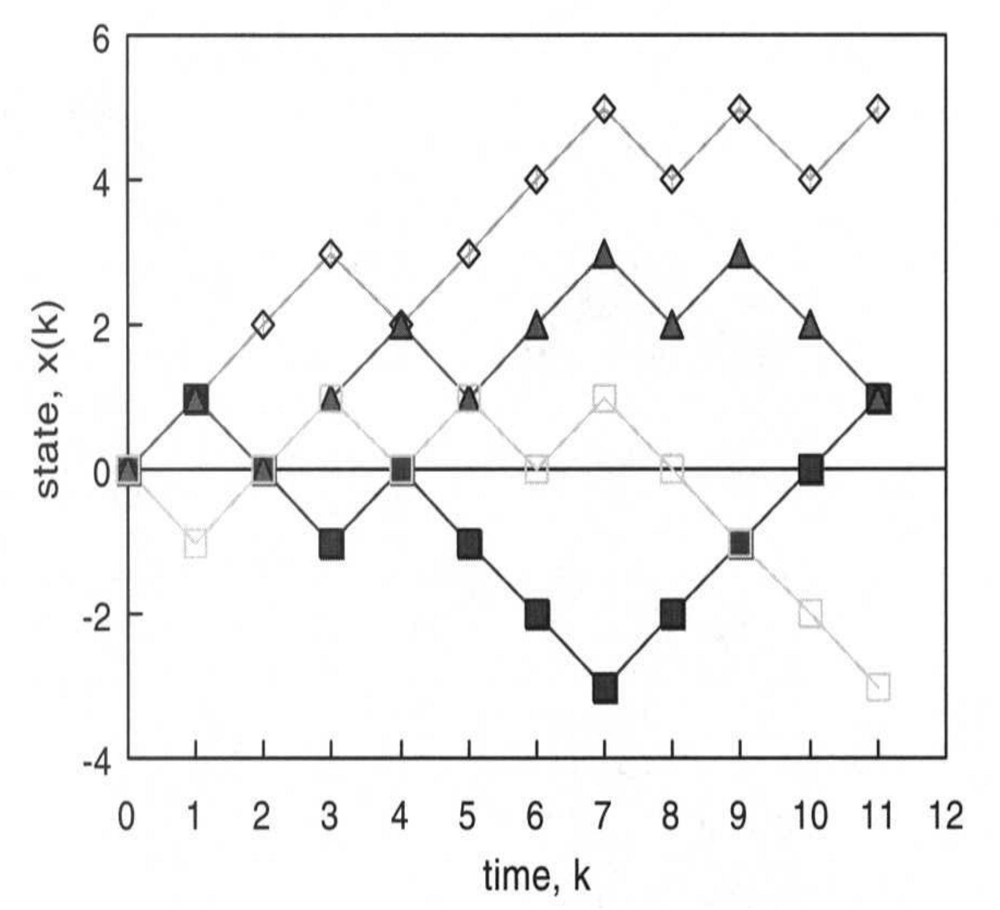
gibi çalışırlar.

RANDOM WALKS

RARTGELEYÜRÜYÜŞLER

RASTGELE YÜRÜYÜŞLER

- Rastgele sürecin özel bir halidir.
- Her bir zamanda artan veya azalan bir şekil çizer.
- Her bir zaman anında rastgele bir doğrultuda hareket eden bir yürüyüşçüye benzer.


6
4
2
0
-2
-4
state, x(k)
0
1
2
3
4
5
6
7
8
9
10
11
12
time, k

- Yürüyenlerin pozisyon ilerleme olasılığı p ise , gerileme olasılığı $1 - p$'dir.
- Yürüyen $x = 0$ konumunda ise, bir sonraki izin verilen pozisyonlar, $x = 1$ ve $x = -1$'dir.
- Negatif ve pozitif yönde sınır olmadığı varsayılırsa, sistemin olasılığı

$$\Pr[X(k) = n]$$

$$= \binom{k}{\frac{1}{2}(k+n)} p^{(k+n)/2} (1-p)^{\frac{k-n}{2}}$$

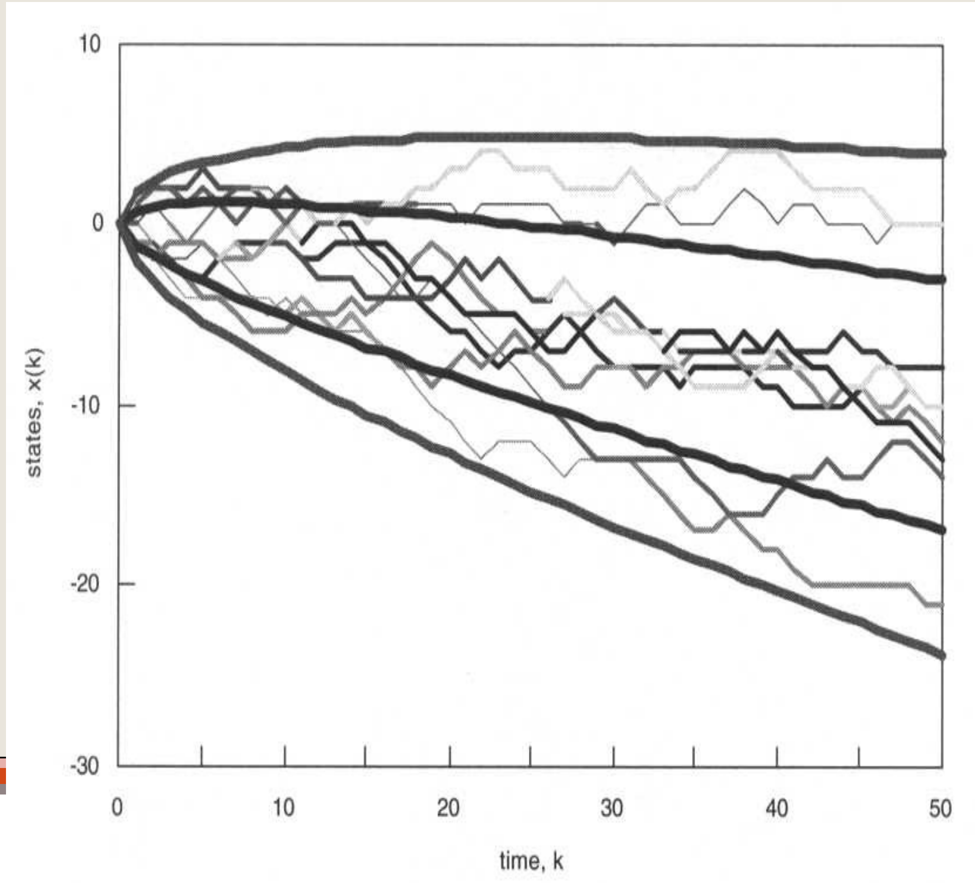
$$n = -k, -k+2, \dots, k-2, k \qquad için.$$

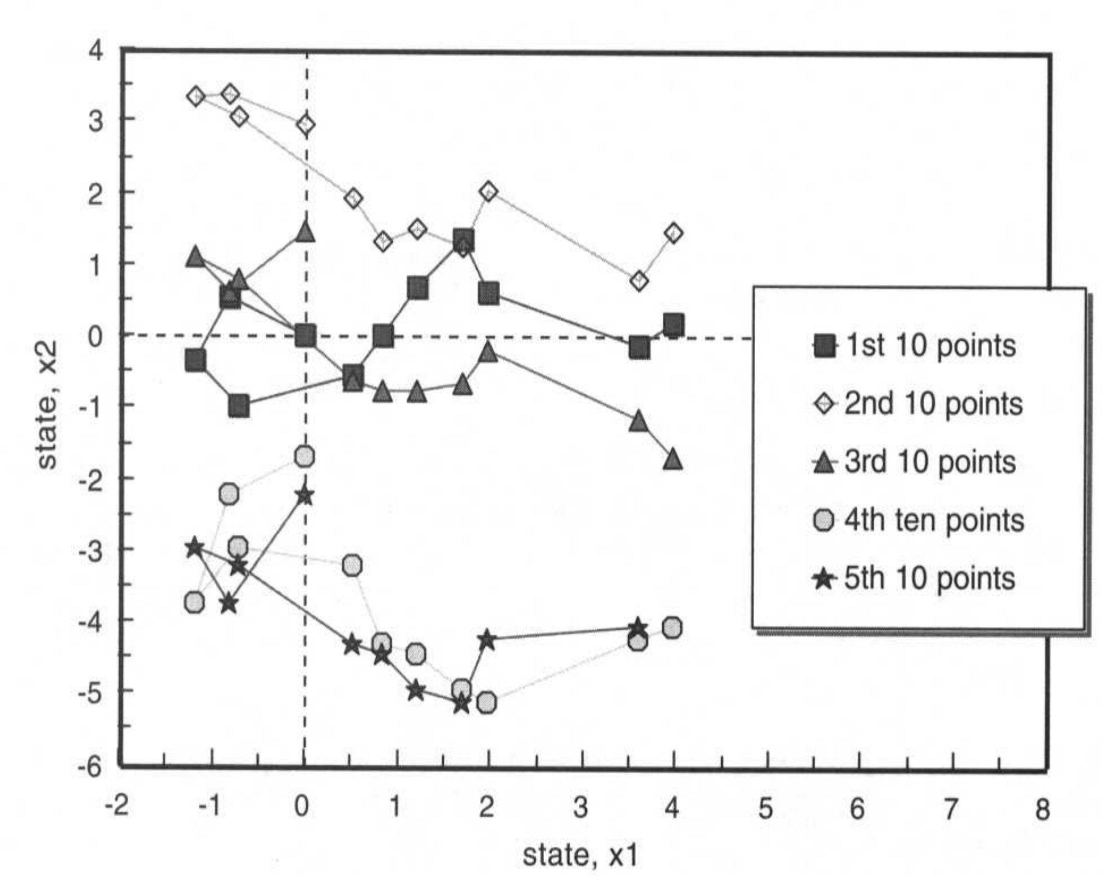
- Sistemin ortalaması;

$$\mu_X(k) = k(2p-1)$$

- Sistemin standart sapması;

$$\sigma^2 = 4kp(1-p)$$


10
0
-10
-20
-30
states, x(k)
0
10
20
30
40
50
time, k



state, x2
state, x1
1st 10 points
2nd 10 points
3rd 10 points
4th ten points
5th 10 points

Algoritma

$$x_1 = 0$$
$$x_2 = 0$$

$$for\ \ k = 1\ \ to\ \ n$$

$$x_1 = \ x_1 + \sqrt{-2 * ln(RND)} * cos(2\pi * RND)$$

$$x_2 = \ x_2 + \sqrt{-2 * ln(RND)} \ * cos(2\pi * RND)$$

$$print\ x_1, x_2$$

$$next\ \ k$$

Adım boyutları Gauss rastgele değişkeni belirlenen rastgele yürüyüş algoritması

WHITE NOISE

BEYAZ GÜRÜLTÜ

BEYAZ GÜRÜLTÜ

- Gürültü modellemek önceden bilinmeyen bir sinyali modellemek olduğundan, gürültünün başka bir $X(k)$ sinyaline bağlı olmadığını düşünebiliriz.

- Bir gürültü sinyali $W(k)$'yı ele alalım.
- Çapraz bağıntısı;

$$R_{wx}(\tau) = E[W(k)X(k+\tau)]$$

$$= E[W(k)]E[X(k+\tau)] = \mu_x \mu_w$$

- Otokorelasyonu;

$$R_{ww}(\tau) = \begin{cases} sabit, & \tau = 0, \\ 0, & \tau \neq 0. \end{cases}$$

Sıfır gecikme zamanı olana kadar sinyalin herhangi bir bağıntısı yoktur.

- Ayrık bir sinyal için sabitin değeri, ikinci andır; $R_{ww}(0) = E[W^2(k)]$.

- Gürültü ortalaması $\mu_w = 0$ olan gürültü beyaz gürültü olarak tanımlanır.