Rastgele Sayı Üretimi

Doç. Dr. İlhan AYDIN

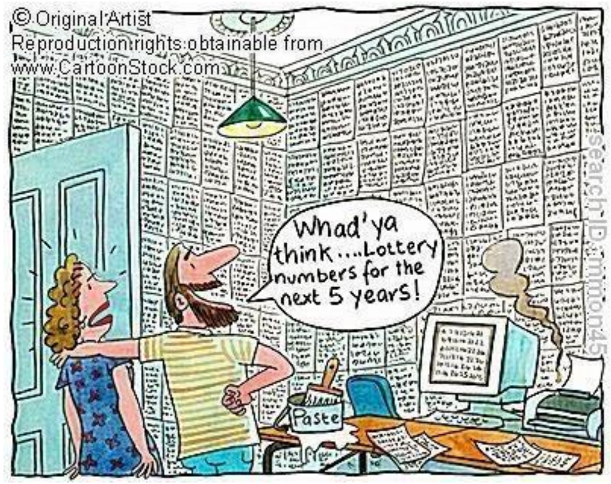
içindekiler

Genel Bakış

- Özellikleri tartışmak ve rastgele sayı üretimi
- Daha sonra, rastgelelik için testlerin tanıtımı :
  - Frekans testi
  - Otokorelasyon testi

Tarihsel olarak

- Tarihsel olarak
  - zar atmak
  - Kart dağıtma
  - Numaralı Top Çiz
  - π rakamlarını kullanın
  - Mekanik cihazlar (dönen disk, vb.)
  - Elektrik Devreleri
    - Elektronik Rastgele Sayı Göstergesi (ERNIE)
    - Gama Işınlarını Saymak
- Bilgisayar ile birlikte
  - Elektronik bir cihazı bilgisayara bağlayın
  - Rasgele sayılar tablosunu okuma



Sözde Rastgele sayılar

Sözde Rassal Sayılar

- Yaklaşım: Aritmetik üretim (hesaplama)rastgele sayılar
- “Sözde”, çünkü bilinen bir numara kullanarak yöntemi gerçek rastgelelik potansiyelini ortadan kaldırır.

> *Aritmetik yöntemleri göz önüne alan rastgele rakamlar üretmek elbette ki zor. Çünkü, birkaç kez belirtildiği gibi, rastgele bir sayı diye bir şey yoktur – orada sadece rastgele sayılar üretmek için kullanılan yöntemlerdir ve katı bir aritmetik işlem elbette böyle bir yöntem değildir.*
>
> *John von Neumann, 1951*

Sözde Rastgele Sayılar

> *... Muhtemelen... haklı gösterilemez, sadece sonuçlarıyla değerlendirilmelidir. Belirli bir tarifle oluşturulan rakamlarla ilgili bazı istatistiksel çalışmalar yapılmalıdır, ancak kapsamlı testler pratik değildir. Rakamlar bir problemde iyi çalışıyorsa, genellikle aynı türden başkalarıyla başarılı olurlar.*
>
> John von Neumann, 1951

- Amaç: [0,1] 'de rasgele sayıların (RN) ideal özelliklerini simüle eden veya taklit eden bir sayı dizisi üretmek.

Sözde Rastgele Sayılar

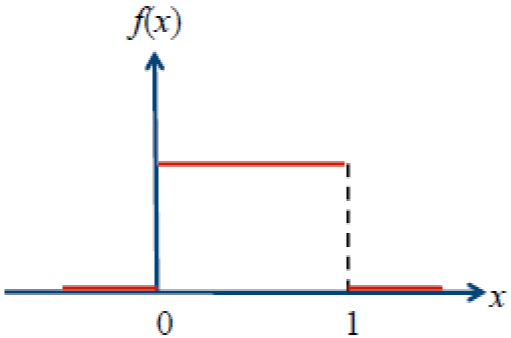
- İyi rasgele sayı rutinlerinin önemli özellikleri
  - Hızlı
  - Farklı bilgisayarlara taşınabilir
  - Yeterince uzun peryoda sahip olmak
  - Tekrarlanabilir
  - Doğrulama ve hata ayıklama
- Farklı sistemler için aynı rasgele sayı akışını kullanın
  - Aşağıdaki istatistiksel özellikleri sağlamalı
    - tekdüzelik
    - bağımsızlık

Sözde Rastgele sayılar

- İki Önemli özellik:
  - Tek düzelik
  - Bağımsızlık
- R_i rastgele sayısı bir olasılık dağılım fonksiyonu(Probability distribution function-PDF) ile tek düzelikten bağımsız olmalıdır.

$$f(x) = \begin{cases} 1, & 0 \le x \le 1 \\ 0, & \text{otherwise} \end{cases}$$

$$E(R) = \int_0^1 x\,dx = \left.\frac{x^2}{2}\right|_0^1 = \frac{1}{2}$$



PDF for random numbers

Sözde Rastgele Sayılar

- Sözde rasgele sayılar üretilirken yaşanan sorunlar
  - Oluşturulan sayılar eşit dağılmamış olabilir
  - Oluşturulan sayılar yerine ayrık değerli olabilir-Sürekli değerli
  - Oluşturulan sayıların ortalaması çok yüksek veya çok düşük olabilir
  - Oluşturulan sayıların varyansı çok yüksek veya çok düşük

- Bağımlılık olabilir:
  - Sayılar arasında otokorelasyon
  - Sayılar bitişik sayılara göre art arda daha yüksek veya daha düşük
  - Ortalamanın üzerinde birkaç sayı ve ardından birkaç sayı ortalamanın altındaki sayılar

Rastgele Sayı Üretimi

Generating Random Numbers

- Orta kare yöntemi
- Doğrusal Eşlenik Yöntem (LCG)
- Kombine Lineer Konjügasyon Jeneratörleri (CLCG)
- Rastgele Sayı Akışı

Rastgele Sayı Üretimi

Orta Kare Yöntemi

Orta Kare Metodu

- İlk aritmetik üreteç: orta kare yöntemi
  - 1940'larda von Neumann ve Metropolis
- Orta kare yöntemi:
  - Dört basamaklı pozitif tamsayı Z0 ile başlayın
  - Hesapla:($Z_0^2 = Z_0 \times Z_0$)8 dijitlik bir tamsayı elde etmek için kare al
  - Ortadaki dört rakamı sonraki değerleri üretmek için kullan.

i	Z_i	U_i	$Z_i \times Z_i$
0	7182	-	51581124
1	5811	0.5811	33767721
2	7677	0.7677	58936329
3	9363	0.9363	87665769
…			

Orta Kare Metodu

- Problem: Üretilen sayılar sıfıra doğru gidebilir.

i	Z_i	U_i	$Z_i \times Z_i$
0	7182	-	51581124
1	5811	0,5811	33767721
2	7677	0,7677	58936329
3	9363	0,9363	87665769
4	6657	0,6657	44315649
5	3156	0,3156	09960336
6	9603	0,9603	92217609
7	2176	0,2176	04734976
8	7349	0,7349	54007801
9	78	0,0078	00006084
10	60	0,006	00003600
11	36	0,0036	00001296
12	12	0,0012	00000144
13	1	0,0001	00000001
14	0	0	00000000
15	0	0	00000000

... Rastgele seçilen yöntemle rastgele sayılar üretilmemelidir. Bazı teoriler kullanılmalıdır.

Donald E. Knuth, The Art of Computer Programming, Vol. 2

Rastgele Sayı Üretimi

Doğrusal Eşlenik Yöntem (LCG)

Doğrusal Eşlenik Yöntem (LCG)

- Özyinelemeli bir ilişki izleyerek X1, X2,… 0 ve m-1 arasında bir tamsayı dizisi üretmek için:

$$X_{i+1} = (\underbrace{a}_{\text{Çarpan}} X_i + \underbrace{c}_{\text{Artış}}) \bmod \underbrace{m}_{\text{Mod}}, \quad i = 0,1,2,...$$

- Varsayımlar: $m > 0$ ve $a < m$, $c < m$, $X_0 < m$
- A, c, m ve X0 için değerlerin seçimi istatistiksel özellikleri ve döngü uzunluğunu büyük ölçüde etkiler
- Rastgele Xi tam sayıları [0, m-1] 'de üretilmektedir.

Doğrusal Eşlenik Yöntem

- Tamsayı Xi'yi rasgele sayılara dönüştürün

$$R_i = \frac{X_i}{m}, \quad i = 1,2,...$$

- Not:
  - $X_i \in \{0, 1, ..., m\text{-}1\}$
  - $R_i \in [0, (m\text{-}1)/m]$

Doğrusal Eşlenik Yöntem: Örnek

- $X_0 = 27$, $a = 17$, $c = 43$ ve $m = 100$'ü kullan.
- X_i ve R_i değerleri:

$$X_1 = (17 \times 27 + 43) \bmod 100 = 502 \bmod 100 = 2 \quad \Rightarrow \quad R_1 = 0.02$$
$$X_2 = (17 \times 2 + 43) \bmod 100 = 77 \quad \Rightarrow \quad R_2 = 0.77$$
$$X_3 = (17 \times 77 + 43) \bmod 100 = 52 \quad \Rightarrow \quad R_3 = 0.52$$
$$X_4 = (17 \times 52 + 43) \bmod 100 = 27 \quad \Rightarrow \quad R_3 = 0.27$$

...

Doğrusal Eşlenik Yöntem: Örnek

- $a = 13$, $c = 0$, ve $m = 64$'ü kullan
- Jeneratörün süresi çok düşük
- Çekirdek X_0 diziyi etkiler

i	X_i $X_0=1$	X_i $X_0=2$	X_i $X_0=3$	X_i $X_0=4$
0	1	2	3	4
1	13	26	39	52
2	41	18	59	36
3	21	42	63	20
4	17	34	51	4
5	29	58	23	
6	57	50	43	
7	37	10	47	
8	33	2	35	
9	45		7	
10	9		27	
11	53		31	
12	49		19	
13	61		55	
14	25		11	
15	5		15	
16	1		3	

Doğrusal Eşlenik Yöntemi:
İyi bir üretecin özellikleri

- Maksimum Yoğunluk
  - R_i tarafından kabul edilen değerler, i = 1,2,…

  [0,1] üzerinde büyük boşluk bırakmayın
  - Sorun: Sürekli değil, her Ri ayrıktır
  - Çözüm: m modülü için çok büyük bir tam sayı
    - Yaklaşıklığın çok az sonucu olduğu görülmektedir.

- Maksimum Period
  - Maksimum yoğunluk elde etmek ve periyodik işletmeden kaçınmak için uygun a, c, m, ve X_0 seçimi ile elde edilir.

- Çoğu dijital bilgisayar sayıların ikili gösterimini kullanır
  - Hız ve verimlilik, 2�much nin üssü m değeri (veya ona yakın) seçilerek elde edilir.

Doğrusal Eşlenik Yöntem:
İyi bir Jeneratörün özellikleri

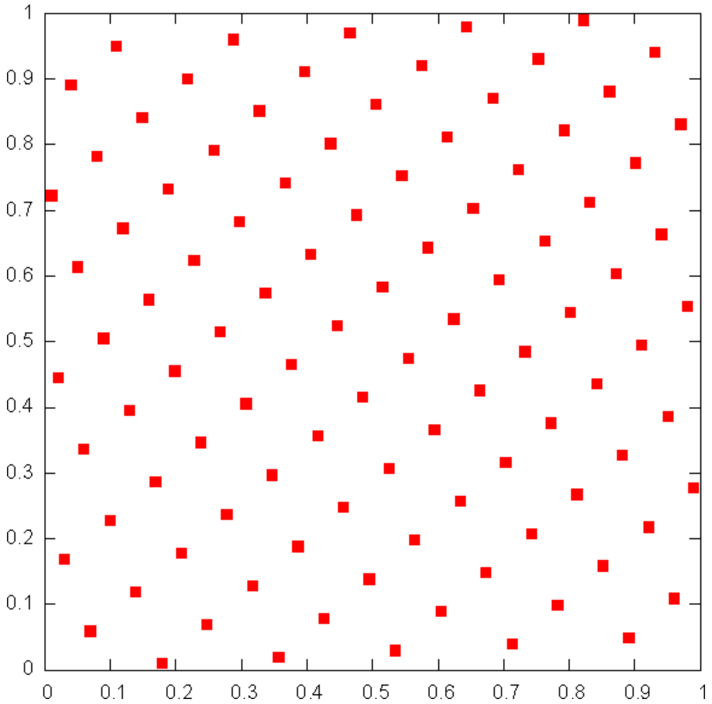
- LCG, ancak aşağıdaki üç koşul geçerliyse tam periyoda sahiptir (Hull ve Dobell, 1962):
  1. Hem m'yi hem de c'yi bölen tek pozitif tamsayı 1'dir.Yani m ve c aralarında asaldır.
  2. q, m'yi bölen asal bir sayıysa, o halde q a-1'i böler
  3. Eğer m 4’e tam böülünebiliyorsa, 4, a-1'i böler

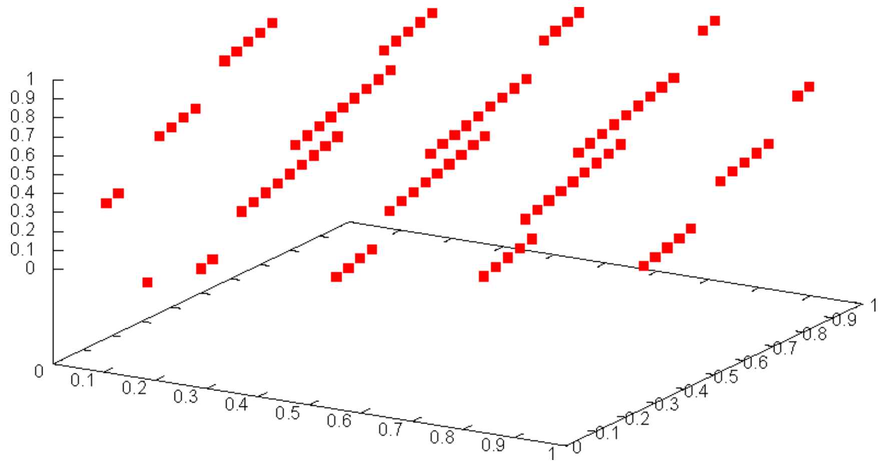
Doğrusal Eşlenik Yöntem:

Uygun parametre seçimi

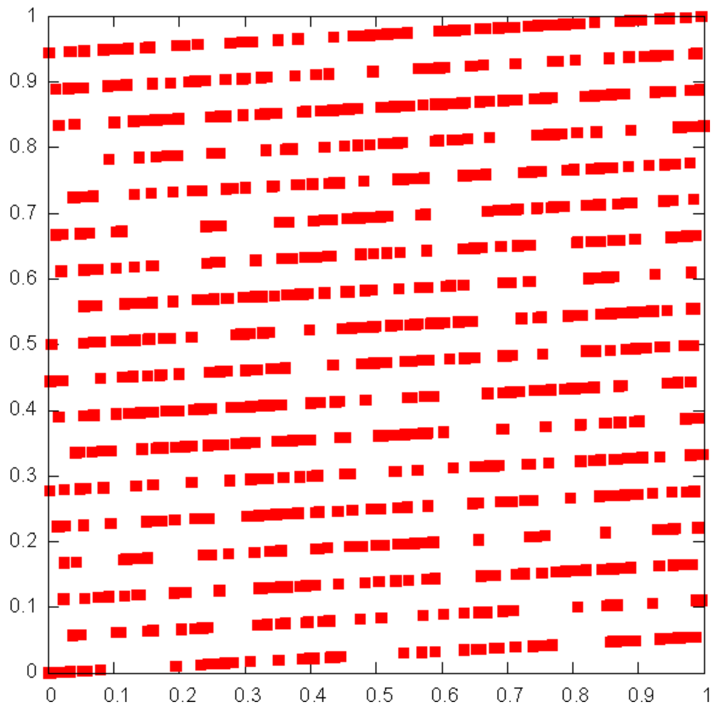
- Eğer m 2'nin kuvveti ise, $m=2^b$, ve $c\neq 0$
  - C ile m aralarında asal ise ve k = bir tamsayı olduğunda $a=1+4k$ ise, mümkün olan en uzun periyod $P=m=2^b$'ye ulaşılır.
- Eğer m 2'nin kuvveti ise, $m=2^b$, ve $c=0$
  - X_0 başlangıcı tek ve $a=3+8k$ veya $a=5+8k$, k = 0,1, için ise mümkün olan en uzun süre $P=m/4=2^{b-2}$ elde edilir ...
- Eğer m asal ve c = 0 ise
  - Mümkün olan en uzun periyod P=m-1 olması için en küçük k değeri için a^k-1 m ile bölünebilmesi gerekir.

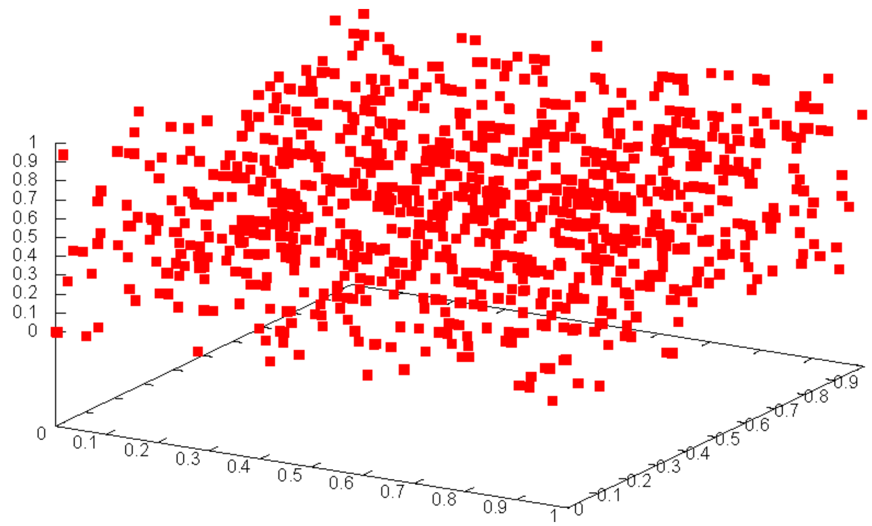
İyi Bir Üretecin Özellikleri





İyi Bir Üretecin Özellikleri





Java'da Rastgele Sayılar

- java.util.Random içinde tanımlanmıştır

```
private final static long multiplier = 0x5DEECE66DL; // 25214903917
private final static long addend = 0xBL;             // 11
private final static long mask = (1L << 48) - 1;     // 2^48-1 = 281474976710655

protected int next(int bits) {
    long oldseed, nextseed;
    ...
    oldseed = seed.get();
    nextseed = (oldseed * multiplier + addend) & mask;
    ...
    return (int)(nextseed >>> (48 - bits));    // >>> Unsigned right shift
}
```

Genel Eşlenik Jeneratörler

- Doğrusal Eşlenik üreteçler, aşağıdakiler tarafından tanımlanan özel bir üreteç örneğidir:

$$X_{i+1} = g(X_i, X_{i-1}, \ldots) \bmod m$$

- burada g () önceki X_i'lerin bir fonksiyonudur.
  - $X_i \in [0, m\text{-}1],\ R_i = X_i/m$

- İkinci dereceden uyumlu üreteç
  - Tanımı: $g(X_i, X_{i-1}) = aX_i^2 + bX_{i-1} + c$
- Çoklu yinelemeli üretçeler
  - Tanımlama: $g(X_i, X_{i-1}, \ldots) = a_1 X_i + a_2 X_{i-1} + \ldots + a_k X_{i-k}$
- Fibonacci üreteci
  - Tanımlayan:

$$g(X_i, X_{i-1}) = X_i + X_{i-1}$$

Kombine Doğrusal Eşlenik Jeneratörler

- Sebep: Simüle edilen sistemlerin karmaşıklığının artması nedeniyle daha uzun periyod üretecine ihtiyaç duyulmaktadır.
- Yaklaşım: İki veya daha fazla çarpan uyumlu üreteci birleştirin.

- $X_{i,1}, X_{i,2}, \ldots, X_{i,k}$ k farklı çarpımsal eşlenik üretecin i'ninci çıkışı olsun.
  - J'ninci jeneratörü $X_{\cdot,j}$:

$$X_{i+1,j} = (a_j X_i + c_j) \bmod m_j$$

    - asal modülü m_j, çarpan a_j ve periyodu m_j -1
    - tamsayı üretir $X_{i,j}$ yaklaşık ~ Düzgün dağılımda $[0, m_j - 1]$
    - $W_{i,j} = X_{i,j} - 1$ *yaklaşık ~ [0, mj - 2] üzerindeki tamsayılarda düzgün (Wi,j = Xi,j - 1 is approx ~Uniform on integers on [0, mj - 2])*

Kombine Doğrusal Eşlenik Jeneratörler

- Önerilen form:

$$X_i = \left(\sum_{j=1}^{k} (-1)^{j-1} X_{i,j} \right) \bmod m_1 - 1 \qquad \text{Hence, } R_i = \begin{cases} \dfrac{X_i}{m_1}, & X_i > 0 \\ \dfrac{m_1 - 1}{m_1}, & X_i = 0 \end{cases}$$

- bundan dolayı maksimum periyod:

$$P = \frac{(m_1 - 1)(m_2 - 1)\ldots(m_k - 1)}{2^{k-1}}$$

Kombine Doğrusal Eşlenik Jeneratörler

- Örnek: 32 bit bilgisayarlar için, k = 2 üreteci $m_1 = 2147483563$, $a_1 = 40014$, $m_2 = 2147483399$ and $a_2 = 40692$ ile birleştirir.
  Algoritma şöyle olur:
  Adım 1: Başlangıç değerlerini seçin
  1. üretec için $[1, 2147483562]$ aralığında $X_{0,1}$
  2. üretec için $[1, 2147483398]$ aralığında $X_{0,2}$
  Adım 2: Her bir üreteç için,
  $X_{i+1,1} = 40014 \times X_{i,1} \mod 2147483563$
  $X_{i+1,2} = 40692 \times X_{i,2} \mod 2147483399$
  Adım 3: $X_{i+1} = (X_{i+1,1} - X_{i+1,2}) \mod 2147483562$
  Adım 4: return

$$R_{i+1} = \begin{cases} \dfrac{X_{i+1}}{2147483563}, & X_{i+1} > 0 \\ \dfrac{2147483562}{2147483563}, & X_{i+1} = 0 \end{cases}$$

  Adım5: i = i + 1 olarak ayarlayın, 2. adıma geri dönün.
- Kombine üretecin periyodu: $(m_1 - 1)(m_2 - 1)/2 \sim 2 \text{ x } 10^{18}$

Excel 2003'te Rasgele Sayılar

- Excel 2003 ve 2007'de yeni Rasgele Sayı Üreticisi

$$X, Y, Z \in \{1, \ldots, 30000\}$$
$$X = X \cdot 171 \bmod 30269$$
$$Y = Y \cdot 172 \bmod 30307$$
$$Z = Z \cdot 170 \bmod 30323$$

$$R = \left(\frac{X}{30269} + \frac{Y}{30307} + \frac{Z}{30323} \right) \bmod 1.0$$

- Bu yöntemin 10^{13} 'den fazla sayı ürettiği belirtilmektedir
- Daha fazla bilgi için: http://support.microsoft.com/kb/828795

Rasgele Sayı Akışı

- Doğrusal bir eşlenik rasgele sayı üreteci için başlangıç :
  - X0 tam sayı değeri rastgele bir sayı serisi ile başlatılıyor mu?
  - Dizideki herhangi bir değer $(X_0, X_1, \ldots, X_P)$) üreteci “başlangıç” için kullanılabilir

- Rasgele sayı akışı:
  - Seriden $(X_0, X_1, \ldots, X_P)$ alınan bir başlangıç değeri seçimi.
  - Akışların birbirinden ayrı b değerleri olması durumunda, i akışı başlangıç tanımlayarak oluştururlur:

$$S_i = X_{b(i-1)} \qquad i = 1, 2, \ldots, \left\lfloor \frac{P}{b} \right\rfloor$$

  - Önceki üreteçler: $b = 10^5$
  - Daha yeni üreteçler: $b = 10^{37}$

- K akışlı tek bir rasgele sayı üreteci k farklı sanal rasgele sayı üreteci gibi davranabilir
- İki veya daha fazla alternatif sistemi karşılaştırmak.
  - Sözde rasgele sayı dizisinin bölümlerini simüle edilen sistemlerin her birinde aynı amaca ayırmak avantajlıdır.

Rastgele Sayı Testleri

Rastgele Sayı Testleri

- İki kategori:
  - **Tekdüzelik** testi:

    $$H_0: \quad R_i \sim U[0,1]$$
    $$H_1: \quad R_i \nsim U[0,1]$$

    - Sıfır hipotezinin H_0 reddedilmemesi, non-uniform kanıtının tespit edilmediği anlamına gelir.
  - **Bağımsızlık** testi:

    $$H_0: \quad R_i \sim \text{bağımsız}$$
    $$H_1: \quad R_i \nsim \text{bağımsız}$$

    - Sıfır hipotezinin , H_0, reddedilmemesi, bağımlılık kanıtı tespit edilmediği anlamına gelir.

- Anlamlılık seviyesi α, Doğru olduğunda H_0'ı reddetme olasılığı:

$$\alpha = P(\text{reject } H_0 \mid H_0 \text{ is true})$$

Rastgele Sayı Testleri

- Bu testler ne zaman kullanılır:
  - İyi bilinen bir simülasyon dili veya rasgele sayı üreteci kullanılıyorsa, muhtemelen test etmek gereksizdir
  - Üreteç açık bir şekilde bilinmiyorsa veya belgelenmiyorsa, örneğin elektronik tablo programları, sembolik / sayısal hesap makineleri, testler birçok örnek numarasına uygulanmalıdır.

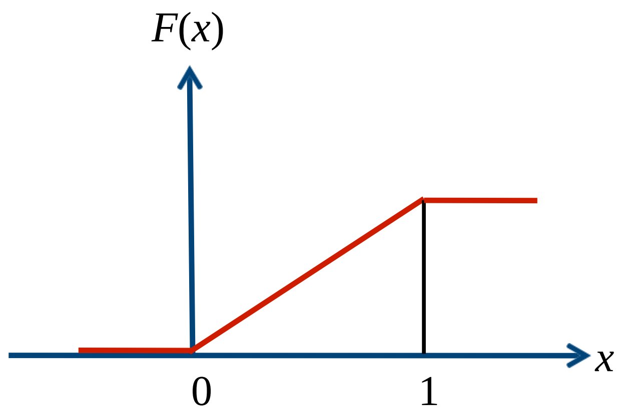
- Test türleri:
  - Teorik testler: Gerçekten sayı üretmeden m, a, ve c seçeneklerini değerlendirin
  - Ampirik testler: Üretilen gerçek sayı dizilerine uygulanır.
    - Bizim amacımız.

Rastgele Sayı Testleri

Frekans testleri: Kolmogorov-Smirnov Testi

Kolmogorov-Smirnov Test

- Düzgün dağılımın sürekli CDF, $F(x)$, N örnek gözlemlerinin ampirik CDF, $S_N(x)$ ile karşılaştırılır.
  - Bilinen durum: $F(x) = x, \quad 0 \le x \le 1$
  - RNG'den alınan örnek $R_1, R_2, \ldots, R_N$ ise, ampirik CDF, $S_N(x)$:



$$S_N(x) = \frac{R_i \le x \quad iken\ Ri\ sayısı}{N}$$

- İstatistiğe dayanarak: $D = max\,|\,F(x) - S_N(x)|$
  - D'nin örnekleme dağılımı bilinmektedir

Kolmogorov-Smirnov Test

- Test aşağıdaki adımlardan oluşur
  - **1. Adım: Verileri küçükten büyüğe sıralayın**
    $R_{(1)} \leq R_{(2)} \leq \ldots \leq R_{(N)}$
  - **2. Adım: Hesaplama**

$$D^+ = \max_{1 \leq i \leq N} \left\{ \frac{i}{N} - R_{(i)} \right\}$$

$$D^- = \max_{1 \leq i \leq N} \left\{ R_{(i)} - \frac{i-1}{N} \right\}$$

  - **3. Adım: Hesaplama**
    $D = \max(D^+, D^-)$
  - **4. Adım:** α **önem düzeyi için** D_α**'yi alın**
  - **Adım 5:** $D \leq D_\alpha$ **kabul ederse, aksi takdirde** H_0**'ı reddedin**

Kolmogorov-Smirnov Kritik Değerleri

Degrees of Freedom (N)	$D_{0.10}$	$D_{0.05}$	$D_{0.01}$
1	0.950	0.975	0.995
2	0.776	0.842	0.929
3	0.642	0.708	0.828
4	0.564	0.624	0.733
5	0.510	$\boxed{0.565}$	0.669
6	0.470	0.521	0.618
7	0.438	0.486	0.577
8	0.411	0.457	0.543
9	0.388	0.432	0.514
10	0.368	0.410	0.490
11	0.352	0.391	0.468
12	0.338	0.375	0.450
13	0.325	0.361	0.433
14	0.314	0.349	0.418
15	0.304	0.338	0.404
16	0.295	0.328	0.392
17	0.286	0.318	0.381
18	0.278	0.309	0.371
19	0.272	0.301	0.363
20	0.264	0.294	0.356
25	0.24	0.27	0.32
30	0.22	0.24	0.29
35	0.21	0.23	0.27
Over 35	$\frac{1.22}{\sqrt{N}}$	$\frac{1.36}{\sqrt{N}}$	$\frac{1.63}{\sqrt{N}}$

Kolmogorov-Smirnov Test

- Örnek: Diyelim ki $N=5$ sayı: 0.44, 0.81, 0.14, 0.05, 0.93.

Adım 1:

Adım 2:

i	1	2	3	4	5
$R_{(i)}$	0.05	0.14	0.44	0.81	0.93
i/N	0.20	0.40	0.60	0.80	1.00
$i/N - R_{(i)}$	0.15	0.26	0.16	-	0.07
$R_{(i)} - (i-1)/N$	0.05	-	0.04	0.21	0.13

$R_{(i)}$'yi en küçükten en büyüğe sıralayınt

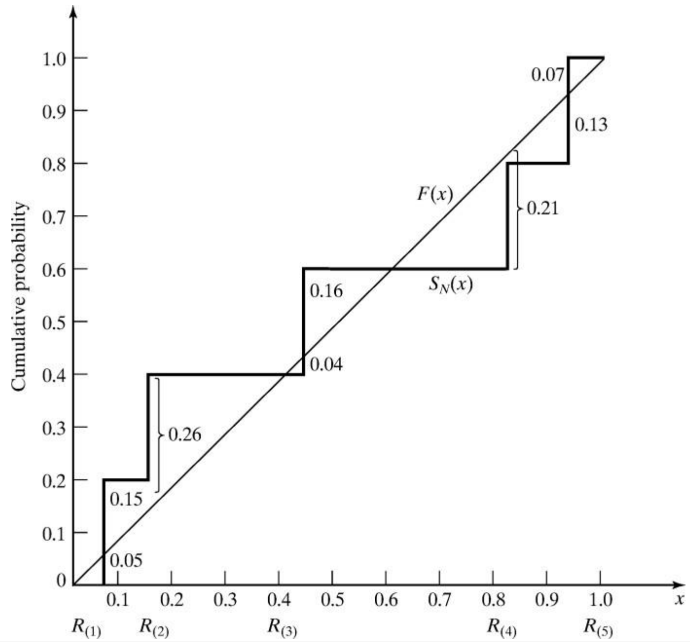
$D^{+} = max\{i/N - R_{(i)}\}$

$D^{-} = max\{R_{(i)} - (i-1)/N\}$

Adım 3: $D = \max(D^{+}, D^{-}) = 0.26$

Adım 4: $\alpha = 0.05$ için,

$D_{\alpha} = 0.565 > D = 0.26$



Bu nedenle, H_0 reddedilmez.

Rastgele Sayı Testleri

Frekans testleri: Chi-kare Testi

Chi-kare (Ki-kare) Test

- Chi-kare testi örnek istatistiği kullanır:

n, sınıfların sayısıdır

O_i *is the observed # in the i-th class*

(O_i i-th sınıfında gözlemlenen #)

E_i *is the expected # in the i-th class*

(Ei i'ninci sınıfında beklenen sayı)

$$\chi_0^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$

  - n-1 serbestlik dereceli yaklaşık ki-kare dağılımı
  - Eşit dağılım için, E_i, her sınıfta beklenen sayı:

$$E_i = \frac{N}{n}, \quad \text{burada N toplam gözlem sayısıdır}$$

- Yalnızca büyük örnekler için geçerlidir, ör. $N \geq 50$

Chi-square Test: Örnek

- [0,1] 'den 100 numaralı örnek, α=0.05
- 10 aralık
- $\chi^2_{0.05,9} = 16.9$
- Kabul et, çünkü
  - $X^2_0 = 11.2 < \chi^2_{0.05,9}$

Interval	Upper Limit	O_i	E_i	O_i-E_i	(O_i-E_i)^2	(O_i-E_i)^2/E_i
1	0.1	10	10	0	0	0
2	0.2	9	10	-1	1	0.1
3	0.3	5	10	-5	25	2.5
4	0.4	6	10	-4	16	1.6
5	0.5	16	10	6	36	3.6
6	0.6	13	10	3	9	0.9
7	0.7	10	10	0	0	0
8	0.8	7	10	-3	9	0.9
9	0.9	10	10	0	0	0
10	1.0	14	10	4	16	1.6
Sum		**100**	**100**	**0**	**0**	**11.2**

$$\chi_0^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$

X^2_0=11.2

Rastgele Sayı Testleri

Otokorelasyon testleri

Otokorelasyon testleri

- Otokorelasyon bir serideki sayılar arasındaki bağımlılıkla ilgilidir
- Örneğin:

0.12	**0.01**	**0.23**	**0.28**	**0.89**	**0.31**	**0.64**	**0.28**	**0.83**	**0.93**
0.99	**0.15**	**0.33**	**0.35**	**0.91**	**0.41**	**0.60**	**0.27**	**0.75**	**0.88**
0.68	**0.49**	**0.05**	**0.43**	**0.95**	**0.58**	**0.19**	**0.36**	**0.69**	**0.87**

- 5., 10., 15., ... sayılar birbirine çok benzer
- Sayılar aşağdaki gibi olabilir.
  - Düşük
  - Yüksek
  - Değişen

Otokorelasyon Testleri

- Her m sayısı arasındaki otokorelasyonun test edilmesi (m gecikmedir), i-inci numaradan başlayarak
  - Sayılar arasındaki otokorelasyon $\rho_{i,m}$: $R_i, R_{i+m}, R_{i+2m}, R_{i+(M+1)m}$
  - M yandaki şartıı sağlayan en büyük tam sayıdır $i + (M+1)m \leq N$
- Hipotez :

$$H_0 : \rho_{i,m} = 0, \quad \text{sayılar bağımsızsa}$$

$$H_1 : \rho_{i,m} \neq 0, \quad \text{eğer sayılar bağımlıysa}$$

- Değerler ilişkisizse:
  - Büyük M değerleri için, $\hat{\rho}_{i,m}$ ile belirtilen $\rho_{i,m}$ tahmincisinin dağılımı yaklaşık olarak normaldir.

Otokorelasyon Testleri

- gecikme j ‘deki korelasyon

$$\rho_j = \frac{C_j}{C_0}$$

$$C_j = Cov(X_i, X_{i+j}) = E(X_i X_{i+j}) - E(X_i)E(X_{i+j})$$

$$C_0 = Cov(X_i, X_i) = E(X_i X_i) - E(X_i)E(X_i) = E(X_i^2) - [E(X_i)]^2 = Var(X_i)$$

$$\Rightarrow \rho_j = \frac{E(X_i X_{i+j}) - E(X_i)E(X_{i+j})}{Var(X_i)}$$

- Varsayalımki $X_i = U_i$

$$E(U_i) = \frac{1}{2} \quad \text{ve} \quad Var(U_i) = \frac{1}{12}$$

$$\rho_j = \frac{E(U_i U_{i+j}) - \frac{1}{4}}{\frac{1}{12}} = 12E(U_i U_{i+j}) - 3$$

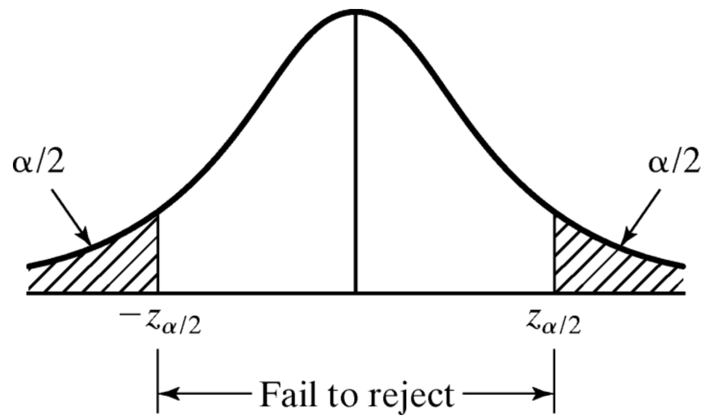
Otokorelasyon Testleri

- Test istatistikleri: $$Z_0 = \frac{\hat{\rho}_{i,m}}{\hat{\sigma}_{\hat{\rho}_{i,m}}}$$
  - Z_0 normalde, ortalama = 0 ve varyans = 1 ile dağıtılır.

$$\hat{\rho}_{i,m} = \frac{1}{M+1}\left[\sum_{k=0}^{M} R_{i+km} \times R_{i+(k+1)m}\right] - 0.25$$

$$\hat{\sigma}_{\rho_{i,m}} = \frac{\sqrt{13M+7}}{12(M+1)}$$

- Z_0 hesaplandıktan sonra $-z_{\alpha/2} \le Z_0 \le -z_{\alpha/2}$ ise bağımsızlık hipotezini reddetmeyin



Otokorelasyon Testleri

- $\rho_{i,m} > 0$ ise, alt sekans pozitif otokorelasyona sahiptir
  - Yüksek rastgele sayıları yüksek sayılar izler ve bunun tersi de geçerlidir.

- $\rho_{i,m} < 0$ ise, alt dizinin negatif otokorelasyonu vardır
  - Düşük rastgele sayıları yüksek sayılar izler ve bunun tersi de geçerlidir.

Örnek

- 38. Slayttaki sayılar için *3rd, 8th, 13th*, ve benzerlerini test edin.
  - Bundan dolayı, $\alpha = 0.05$, $i = 3$, $m = 5$, $N = 30$, ve $M = 4$

$$\hat{\rho}_{35} = \frac{1}{4+1}\begin{bmatrix}(0.23)(0.28)+(0.28)(0.33)+(0.33)(0.27) \\ +(0.27)(0.05)+(0.05)(0.36)\end{bmatrix} - 0.25$$

$$= -0.1945$$

$$\sigma_{\hat{\rho}_{35}} = \frac{\sqrt{13(4)+7}}{12(4+1)} = 0.128$$

$$Z_0 = -\frac{0.1945}{0.1280} = -1.516$$

- $z_{0.025} = 1.96$
- $-1.96 \le Z_0 = -1.516 \le 1.96$ olduğundan, hipotez reddedilmez.

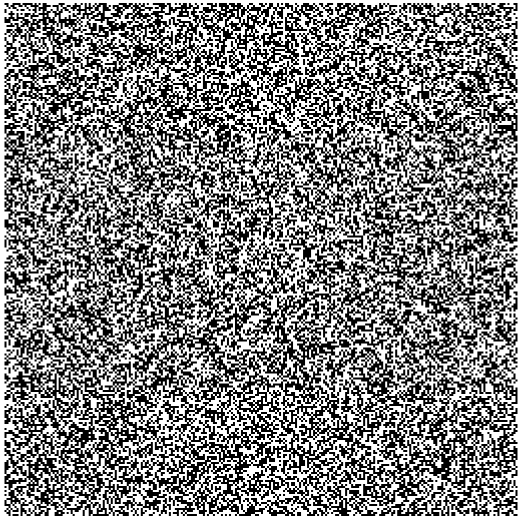
Eksiklikleri

- Test, özellikle test edilen sayılar düşük tarafta olduğunda, küçük M değerleri için çok hassas değildir.
- Çok sayıda test yaparak otokorelasyon için "balık tutma" problemi:
  - $\alpha = 0.05$ ise, 0.05'in gerçek bir hipotezi reddetme olasılığı vardır.
  - 10 bağımsızlık dizisi incelenirse:
    - Sadece tesadüfen önemli bir otokorelasyon bulamama olasılığı 0.95^{10} = 0.60'dır.
    - Bu nedenle, mevcut olmadığında önemli otokorelasyon saptama olasılığı =% 40

Gerçek Rastgele Sayılar

Gerçek Rastgele Sayılar

- İnternette gerçek rastgele sayılar için de kaynaklar var
- www.random.org „RANDOM.ORG internetteki herkese gerçek rastgele sayılar sunuyor. Rasgelelik, birçok amaç için tipik olarak bilgisayar programlarında kullanılan sahte rasgele sayı algoritmalarından daha iyi olan atmosferik gürültüden gelir. İnsanlar sayıları piyango, çekiliş ve çekilişler yapmak, oyunları ve kumar siteleri için kullanıyor. ”



http://www.random.org/analysis/

Gerçek Rastgele Sayılar

- http://www.randomnumbers.info/
„Talep üzerine kuantum rasgele sayı üreteci kullanılarak oluşturulan gerçek rasgele sayıları indirme imkanı sunar. “

Gerçek Rastgele Sayılar

- Donanım tabanlı rasgele sayı üretimi
- http://www.comscire.com

Özet

- Bu bölümde şunları açıkladık:
  - Rasgele sayı üretimi
  - Tekdüzelik ve bağımsızlık testi
  - Gerçek rasgele sayı kaynakları

- Dikkat:
  - Bazıları hala kullanımda olan ve yıllardır kullanılan jeneratörlerde bile yetersizdir.
  - Bu bölüm yalnızca temel bilgileri sunar
  - Ayrıca, üretilen sayılar tüm testleri geçse bile, altta yatan bazı kalıplar tespit edilmemiş olabilir.