# SİMULASYON

*Rastsallık ve Stokastik Süreçler*

**Prof. Dr. H. Kemal İlter**

@hkilter

Tanım

### Rastsallık

*Rastsallık*, düzensizlik durumunu anlatır. Rastsal süreç ise, sonuçları deterministik olarak açıklanamayan ancak bir olasılık dağılımını takip eden olaylar bütünüdür. İstatistiğin konuları içinde sıklıkla karşılaşılan rastsallık, belirli bir korelasyondan (ilişkiden) yoksun olayları istatistiksel olarak açıklamak için kullanılır.

### Yapay-rastsallık

Rastsallık matematiksel bir algoritmayla ortaya çıkarıldıysa *yapay-rastsallık (pseudo-randomness)* olarak adlandırılır. Rastsallık gerçekte öngörülemez olmasına rağmen yapay-rastsallığa sahip bir olayın, aslında o olayın saf (gerçek) rastsal olmadığını, olayın olasılık veya beklenen değer gibi kavramlarla açıklanabilen bir karakteristiğe sahip olduğu görülür.

### Rastsal Değişken

Yapay-rastsal (pseudo-random) sayı kavramı $[0, 1]$ aralığındaki düzgün dağılıma ilişkin rastsal sayıları belirtir. $[0, 1]$ aralığı dışındaki düzgün dağılımına sahip diğer rastsal sayılar, *rastsal değişkenler* veya *stokastik değişkenler* olarak adlandırılır.

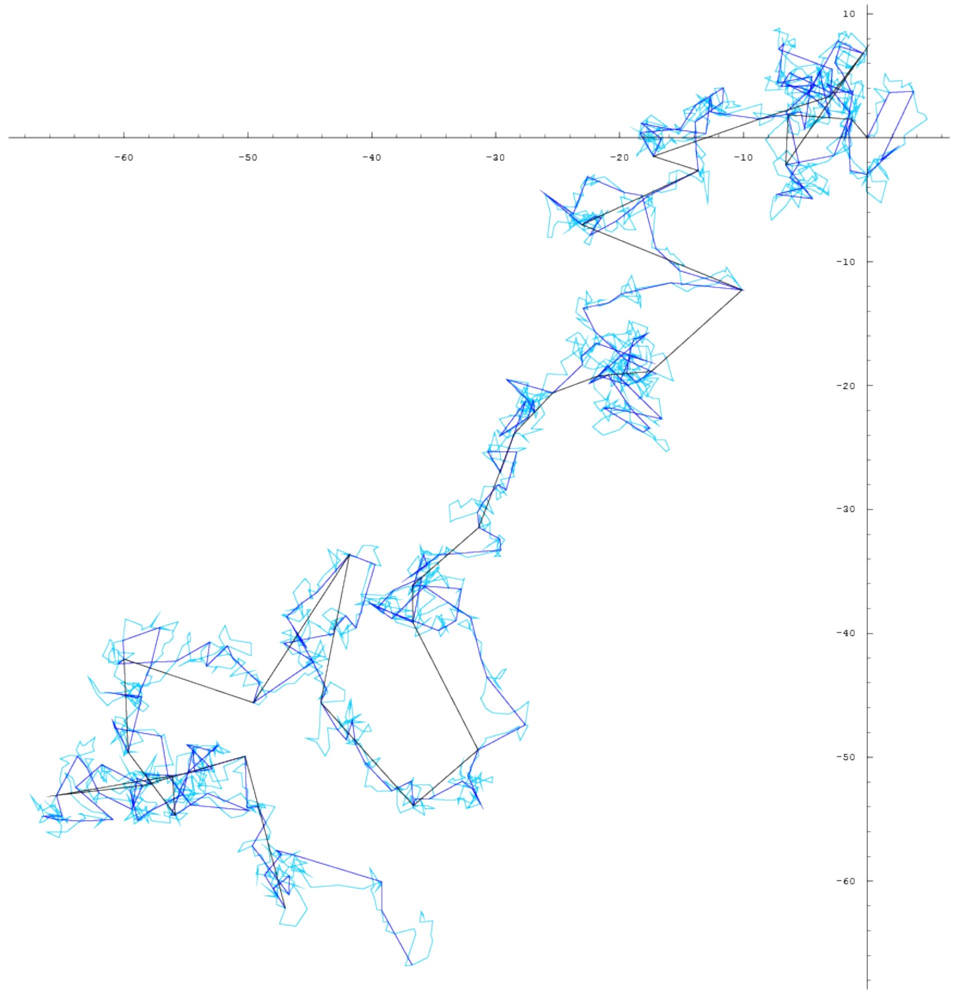
Rastsallık temelde üç durum sonucunda ortaya çıkabilir:

- Rastsallık doğadaki olaylardan kaynaklanıyor olabilir (Örn. Wiener metodu ve Brown hareketi).
- Rastsallık başlangıç durumlarından, diğer bir deyişle başlangıçtaki mikro farklılıkların bir sonucu olarak, oluşuyor olabilir (Örn. Kaos teorisi).
- Rastsallık sistem tarafından üretiliyor olabilir (yapay-rastsallık, Örn. algoritmalar).

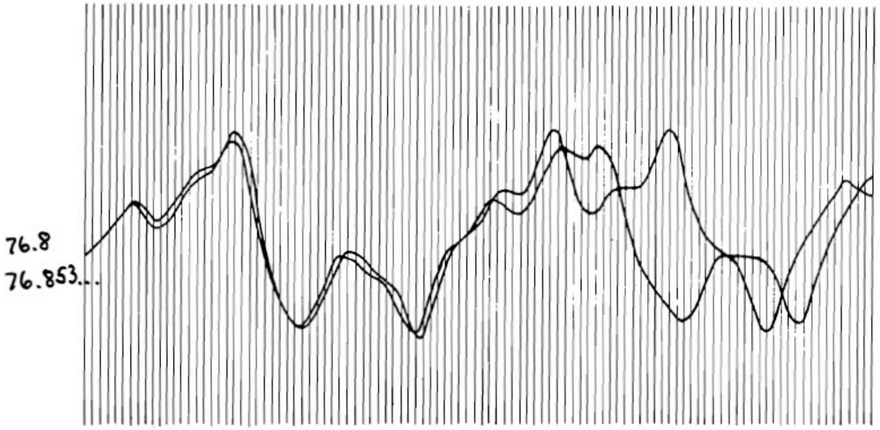
### Weiner Metodu

$t \geq 0$ için $W(0) = 0$ ile sürekli bir stokastik süreç olan $W(t)$'ye ilişkin, $W(t) - W(s)$ artışı, varyansı $0 \leq s < t$ ve ortalaması 0 olan bir Gauss fonksiyonudur. Çakışmayan zaman aralıkları için artışlar bağımsızdır.

### Brown Hareketi



## Oluşum




How two weather patterns diverge. From nearly the same starting point, Edward Lorenz saw his computer weather produce patterns that grew farther and farther apart until all resemblance disappeared. (From Lorenz's 1961 printouts.)

### Deterministic Nonperiodic Flow[1]


Edward N. Lorenz

*Massachusetts Institute of Technology*



(Manuscript received 18 November 1962, in revised form 7 January 1963)



Abstract

Finite systems of deterministic ordinary nonlinear differential equations may be designed to represent forced dissipative hydrodynamic flow. Solutions of these equations can be identified with trajectories in phase space. For those systems with bounded solutions, it is found that nonperiodic solutions are ordinarily unstable with respect to small modifications, so that slightly differing initial states can evolve into considerably different states. Systems with bounded solutions are shown to possess bounded numerical solutions.

A simple system representing cellular convection is solved numerically. All of the solutions are found to be unstable, and almost all of them are nonperiodic.

The feasibility of very-long-range weather prediction is examined in the light of these results.


The stability of a solution $X(\tau)$, $Y(\tau)$, $Z(\tau)$ may be formally investigated by considering the behavior of small superposed perturbations $x_0(\tau)$, $y_0(\tau)$, $z_0(\tau)$. Such perturbations are temporarily governed by the linearized equations

$$\begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix}^{\cdot} = \begin{bmatrix} -\sigma & \sigma & 0 \\ (r-Z) & -1 & -X \\ Y & X & -b \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix}. \qquad (29)$$

Since the coefficients in (29) vary with time, unless the basic state $X$, $Y$, $Z$ is a steady-state solution of (25)–(27), a general solution of (29) is not feasible. However, the variation of the volume $V_0$ of a small region in phase space, as each point in the region is displaced in accordance with (25)–(27), is determined by the diagonal sum of the matrix of coefficients; specifically

$$V_0^{\cdot} = -(\sigma+b+1)V_0. \qquad (30)$$

This is perhaps most readily seen by visualizing the motion in phase space as the flow of a fluid, whose divergence is

$$\frac{\partial X^{\cdot}}{\partial X} + \frac{\partial Y^{\cdot}}{\partial Y} + \frac{\partial Z^{\cdot}}{\partial Z} = -(\sigma+b+1). \qquad (31)$$

Hence each small volume shrinks to zero as $\tau \rightarrow \infty$, at a rate independent of $X$, $Y$, and $Z$. This does not imply that each small volume shrinks to a point; it may simply become flattened into a surface. It follows that the volume of the region initially enclosed by the surface $S$ shrinks to zero at this same rate, so that all trajectories ultimately become confined to a specific subspace having zero volume. This subspace contains all those trajectories which lie entirely within $R$, and so contains all central trajectories.

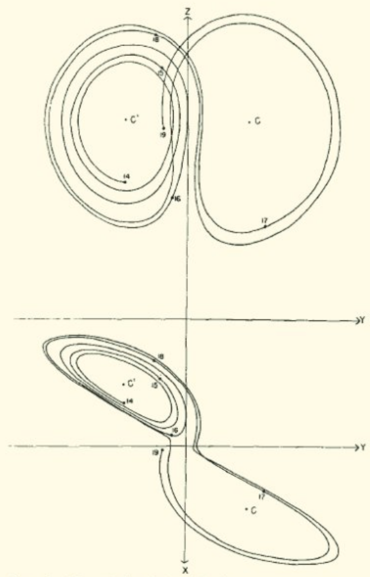


Fig. 2. Numerical solution of the convection equations. Projections on the $X$-$Y$-plane and the $Y$-$Z$-plane in phase space of the segment of the trajectory extending from iteration 1400 to iteration 1900. Numerals "14," "15," etc., denote positions at iterations 1400, 1500, etc. States of steady convection are denoted by $C$ and $C'$.

## Özellikler

**Yöntem**

Rastsal sayılar oluşturmak için kullanılan yöntem, aşağıdaki özelliklere sahip sayı dizileri ortaya çıkarmalıdır:

- düzgün dağılıma (uniform distribution) sahip, (*herhangi bir sayının ortaya çıkma olasılığı diğerleriyle aynı*)
- istatistiksel olarak bağımsız, (*herhangi bir sayı dizideki diğer bir sayının olasılığını etkilemiyor*)
- tekrar üretilebilir, (*belirli bir başlangıç değerine –seed– sahip*)
- dizinin istenen herhangi bir uzunluğu (periyod) için tekrarlanmayan.

**Örnek**

Dizi 1:
0101010101010101010101010101010101010101010101010101010101010101
Dizi 2:
1100100001100001110111101110110011111010010000100101011110010110

*Kolmogorov karmaşıklığı* (Kolmogorov complexity)

Hadamard transform

https://www2.lbl.gov/Science-Articles/Archive/pi-random.html

### Hipotez Testleri

Kendall ve Smith'in orijinal dört rastsallık testi, sıfır hipotezini, $H_0$, belirli bir rastsal dizideki her sayının eşit sayıda olması ve dizideki çeşitli modellerin de eşit olarak dağılması gerektiği fikrine dayanan hipotez testleriydi.

- *Frekans testi*, bir dizideki sayıların (0, 1, 2, ...) kabaca eşit sayıda olup olmadığıyla ilgilidir.
- *Seri test*, frekans testini bir defada iki dijit (00, 01, 02, ...) kontrol edecek şekilde çalıştırarak gözlenen frekansların hipotetik tahminle aynı biçimde dağılması gerektiği fikrine dayanır.
- *Poker testi*, aynı anda beş sayının belirli diziler halinde oluşumunu (AAAAA, AAAAB, AAABB, ...) test etmektedir.
- *Boşluk testi*, 0'lar arasındaki aralığa bakarak bir değerlendirme yapmaktadır. (Örn. 00, 0 uzaklığa; 030, 1 uzaklığa; 02250, 3 uzaklığa sahiptir).

Yerel rastsallık (local randomness)

Gerçek rastsallık (true randomness)

### Yapay-rastsallığa İlişkin Testler

- Monobit
- Wald–Wolfowitz runs test
- Information entropy
- Autocorrelation test
- Kolmogorov–Smirnov test
- Spectral Density Estimation
- Maurer's Universal Statistical Test
- The Diehard tests

https://en.wikipedia.org/wiki/Diehard_tests

## Bazı Yapay-rastsal Sayı Üreteçleri

**Eşleşik üreteç**

$$x_{i+1} = ax_i + c \pmod{m}, \quad x_i, a, c, m \geq 0$$

Örneğin;

$$x_{i+1} = 2x_i + 7 \pmod{9}, \quad x_0 = 1234 \text{ (başlangıç değeri)}$$

ile şu rastsal dizi oluşturulur: $\overline{073461}073461\ldots$

Örneğe ilişkin R kodu aşağıda görülmektedir:

```
# Eşleşik üreteç fonksiyonu
rastsal_dizi <- function(a,c,m,adet,seed) {
    x <- rep(0,adet)
    x[1] <- seed
    for (i in 1:(adet-1)) {
      x[i+1] <- (a * x[i] + c) %% m }
    return(list(x=x))
    }

# Fonksiyonun örnekteki parametreler ile çalıştırılması
z <- rastsal_dizi(2,7,9,12,1234)
z
```

**Tausworthe üreteci**

$$x_i = (a_1 x_{i-1} + a_2 x_{i-2} + \cdots + a_n x_{i-n}) \pmod{2}$$

$$x_i = (1x_{i-1} + 1x_{i-2} + \cdots + 1x_{i-10}) \pmod{2}, \quad x_0 = 1234, a_i = 1, n = 10$$

ile şu rastsal dizi oluşturulur: $\overline{0111111010}1011111110101\ldots$

Örneğe ilişkin R kodu aşağıda görülmektedir:

```
# https://cran.r-project.org/web/packages/SyncRNG/
library(SyncRNG)

# Tausworthe ureteci
rastsal_dizi <- SyncRNG(seed=1234)
  for (i in 1:20) {
    cat(rastsal_dizi$randbelow(2))
  }
```

# RASTSAL SAYI ÜRETMEK

## R

R'da rastsal sayı üretmek için aşağıdaki örnek kullanılabilir, `rnorm` normal dağılımdan rastsal örnekleme yapmak için kullanılmaktadır:

```
# Ortalaması 0 ve standart sapması 1 olan
# 4 adet standart normal rastsal sayı
x <- rnorm(4)
x
# Ortalaması 20 ve standart sapması 2 olan
# 4 adet standart normal rastsal sayı
x <- rnorm(4, 20, 2)
x
# Temel istatistik
summary(x)
```

## Örnek

**Problem**

Bir pastane günlük olarak 20 adet pasta üretmektedir. Pastalara olan günlük talep aşağıdaki olasılık dağılımı ile ortaya çıkmaktadır.

Talebe ait olasılık dağılımı

| *Talep (adet)* | *Olasılık (%)* |
|---|---|
| 5 | 8 |
| 10 | 12 |
| 15 | 25 |
| 20 | 20 |
| 25 | 20 |
| 30 | 15 |

Pastane, pastalar için 15TL birim maliyete katlanmakta ve pastaları 30TL birim fiyatla satmaktadır. Gün içinde satılmayan pasta olursa, onları da gün sonunda yakındaki bir öğrenci yurduna 7.5TL birim fiyatla (maliyetinin altında) vermektedir.

Pasta talebi günlük üretimden fazla ise satış miktarı üretimle sınırlı olacaktır, pasta talebi günlük üretimden az ise satış miktarı talep kadar olacaktır.

Pastanenin 30 günlük pasta satışı ile ilgili simülasyonunu yaparak, ortalama günlük karını hesaplayacağız.