薬学情報処理演習　第5回

表計算ソフトによる統計処理

奥薗　透

コロイド・高分子物性学

データの整理

平均値と分散

- N 個の(数値)データ $x_1, x_2, \cdots, x_N$
- 平均値: $\bar{x} = \frac{x_1+x_2+\cdots+x_N}{N}$
- 分散: $\sigma^2 = \frac{(x_1-\bar{x})^2+(x_2-\bar{x})^2+\cdots+(x_N-\bar{x})^2}{N}$ (σ:標準偏差)

度数分布

- 値$X_1, X_2, \cdots, X_n$ をとる度数(頻度): $F_1, F_2, \cdots, F_n$
- $\bar{x} = \frac{F_1X_1+F_2X_2+\cdots+F_nX_n}{F_1+F_2+\cdots+F_n}$ 　　$N = F_1 + F_2 + \cdots + F_n$
- $\sigma^2 = \frac{F_1(X_1-\bar{x})^2+F_2(X_2-\bar{x})^2+\cdots+F_n(X_N-\bar{x})^2}{F_1+F_2+\cdots+F_n}$

代表値 X_i	度数 F_i
1	1
2	4
3	15
4	9
5	1

チェビシェフの定理

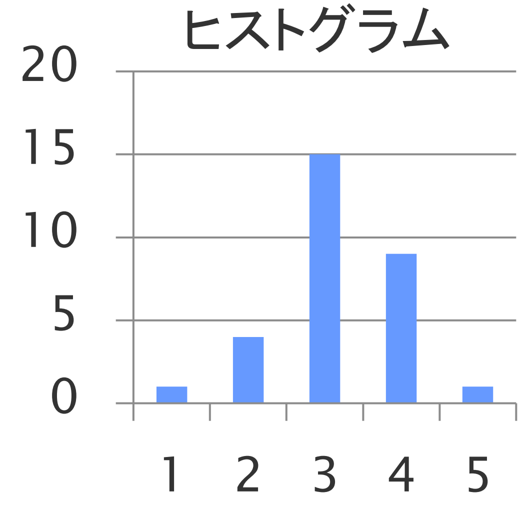
$|x_i - \bar{x}| \le \lambda\sigma \quad (\lambda > 1)$ を満たすデータの個数は $N(1 - 1/\lambda^2)$よりも大きい。

データの分布

- 離散的なデータ
  - 度数分布 F_i
  - 確率 $f_i = F_i/N$

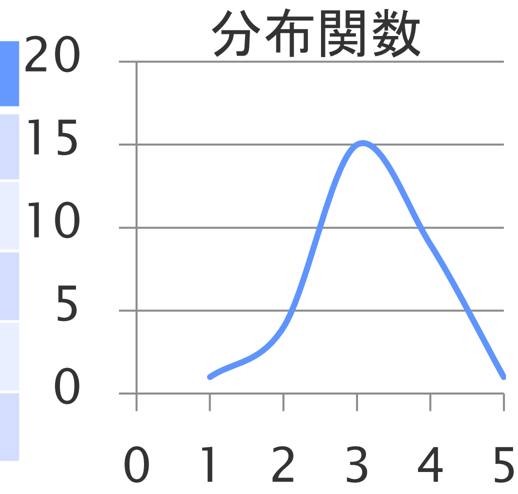
$$\bar{x} = \sum_{i=1}^{n} X_i f_i \qquad \sigma^2 = \sum_{i=1}^{n} (X_i - \bar{x})^2 f_i$$

代表値 X_i	度数 F_i
1	1
2	4
3	15
4	9
5	1



- 連続的なデータ
  - 度数分布→分布関数 $F(x)$
    $(N \to \infty, \Delta x = X_{i+1} - X_i \to 0)$
  - 確率密度 $f(x) = F(x)/\tilde{N}$
    $(\tilde{N} = \int F(x)dx)$

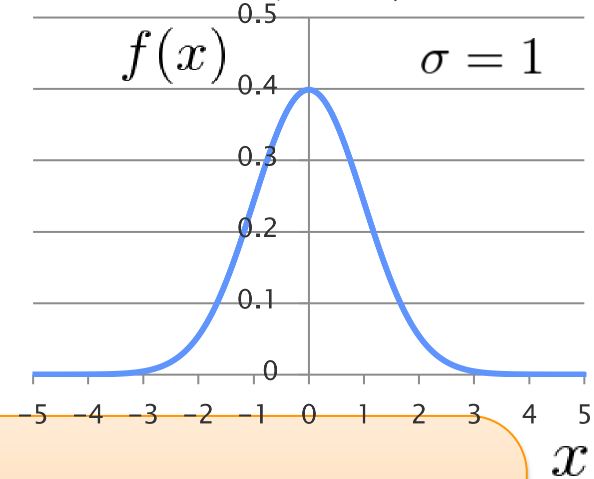
級	X_i	F_i
$0 < x \leq 1$	0.5	1
$1 < x \leq 2$	1.5	4
$2 < x \leq 3$	2.5	15
$3 < x \leq 4$	3.5	9
$4 < x \leq 5$	4.5	1



$$\bar{x} = \int x f(x) dx \qquad \sigma^2 = \int (x - \bar{x})^2 f(x) dx$$

正規分布(ガウス分布)

- 実験的に測定される量には“ばらつき”がある。ばらつき＝平均値からのずれは以下のガウス分布に従うことが多い。なぜか？



$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \mathrm{e}^{-\frac{x^2}{2\sigma^2}}$$

(σ^2 : 分散)

- 中心極限定理

n 個の独立な確率変数 u_i(分散 s_i^2 平均値0)からなる確率変数

$$x_n = (u_1 + u_2 + \cdots + u_n)/\sqrt{\sigma_n^2} \qquad \sigma_n^2 = s_1^2 + s_2^2 + \cdots + s_n^2$$

は、$n \to \infty$ で分散1, 平均値0の正規分布に従う。

- ばらつき＝多数の確率的事象の和

疑似乱数の生成

- コンピュータ上でランダムな数(乱数)を次々に生成し、ランダムなデータを作ることができる。エクセルでは RAND() という関数が用意されている。
- RAND()で生成される乱数は一様分布関数

$$f(x) = 1 \qquad (0 \leq x < 1)$$

に従い、平均と分散は、

$$\overline{x} = \int_0^1 x f(x) dx = \frac{1}{2} \qquad \sigma^2 = \int_0^1 (x - \overline{x})^2 f(x) dx = \frac{1}{12}$$

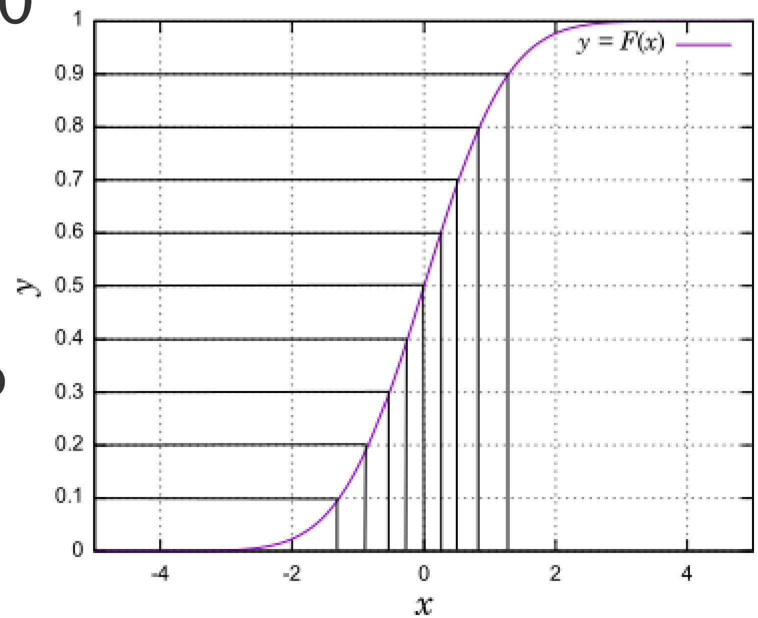
となるので、平均0の一様乱数(RAND()−0.5)を12個足し合わせたものは、近似的に、平均0分散1の正規分布に従う乱数(正規乱数)となっている。

Excel での正規乱数発生方法

- データ分析ツールの利用(関数ではない)
  - *データ/分析/データ分析/乱数発生*
- 中心極限定理の応用
  - (12個の一様乱数の和)−6
  - ＝rand()+rand()+…+rand()−6.0
- 逆関数法
  - ＝norm.inv(rand(),0,1)または＝norm.s.inv(rand())を用いる
  - 分布$f(x)$の累積分布を$F(x)$とする

$$y = F(x) = \int_{-\infty}^{x} f(u)du$$

$$\frac{dy}{dx} = f(x),\ dy = f(x)dx$$



ブラウン運動

- コロイド粒子（微粒子）の乱雑な運動
  - 多数の水分子の衝突の結果
- 1次元のモデル

$$x_{n+1} = x_n + \xi_n \quad (n = 0, 1, 2, \dots)$$

$$\langle \xi_n \xi_m \rangle = \begin{cases} \sigma^2 & (n = m) \\ 0 & (n \neq m) \end{cases}$$

時刻 $t_n = n\Delta t$ での
粒子の位置: x_n
粒子のランダムな変位: ξ_n

$\langle \cdot \rangle$: 平均を表す

- 位置の平均値 $\langle x_n \rangle = 0$
- 平均2乗変位と拡散係数

$$\langle x_n^2 \rangle = \langle (\xi_0 + \xi_1 + \cdots + \xi_{n-1})^2 \rangle$$
$$= \langle \xi_0^2 \rangle + \langle \xi_1^2 \rangle + \cdots + \langle \xi_{n-1}^2 \rangle = n\sigma^2$$

拡散係数:

$$D \equiv \frac{\sigma^2}{2\Delta t} = \frac{\langle x_n^2 \rangle}{2n\Delta t}$$

$$\boxed{x_{n+1} = x_n + \sqrt{2D\Delta t}\mathcal{N}(0,1)}$$

$\langle (aX)^2 \rangle = a^2 \langle X^2 \rangle \quad (\langle X \rangle = 0, \langle X^2 \rangle = 1)$

$\mathcal{N}(0,1)$: 平均0,分散1の正規乱数

ポテンシャル中のブラウン粒子

- ポテンシャル$U(x)$による力を受けながらブラウン運動する質量mの粒子の運動方程式

$$\frac{\mathrm{d}x}{\mathrm{d}t} = v$$

$$m\frac{\mathrm{d}v}{\mathrm{d}t} = -\gamma v - \frac{\mathrm{d}U}{\mathrm{d}x} + \xi$$

γ : 摩擦係数

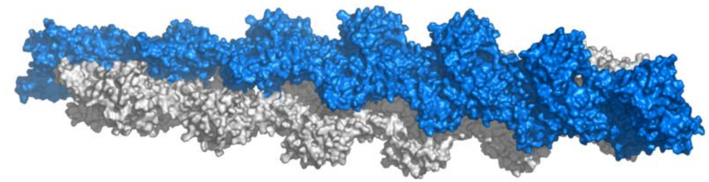
（質量×加速度）＝（摩擦力）＋（ポテンシャル力）＋（熱揺動力）

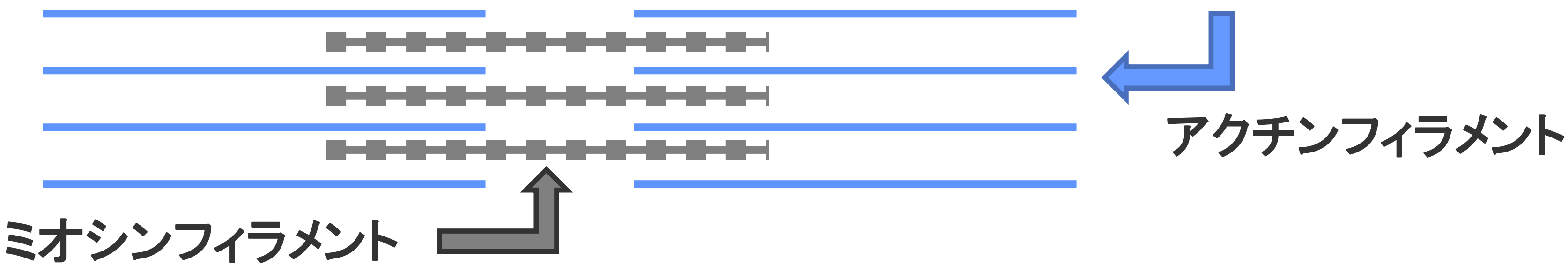
- 加速度が無視できるとき

$$\gamma\frac{\mathrm{d}x}{\mathrm{d}t} = -\frac{\mathrm{d}U}{\mathrm{d}x} + \xi$$

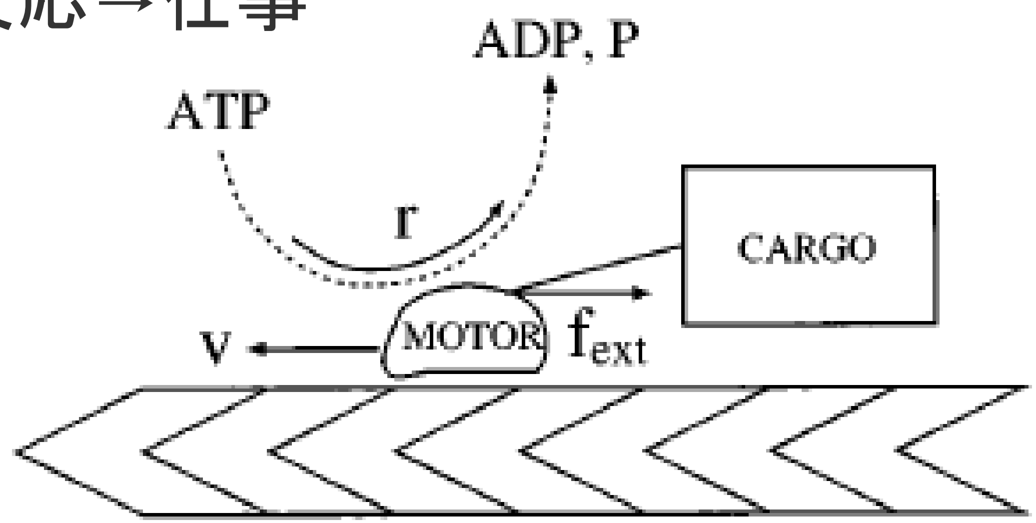
分子モーター

- 筋肉の収縮・弛緩





- 化学反応→仕事



F. Julicher et al., Rev. Mod. Phys. **69**, 1269 (1997)

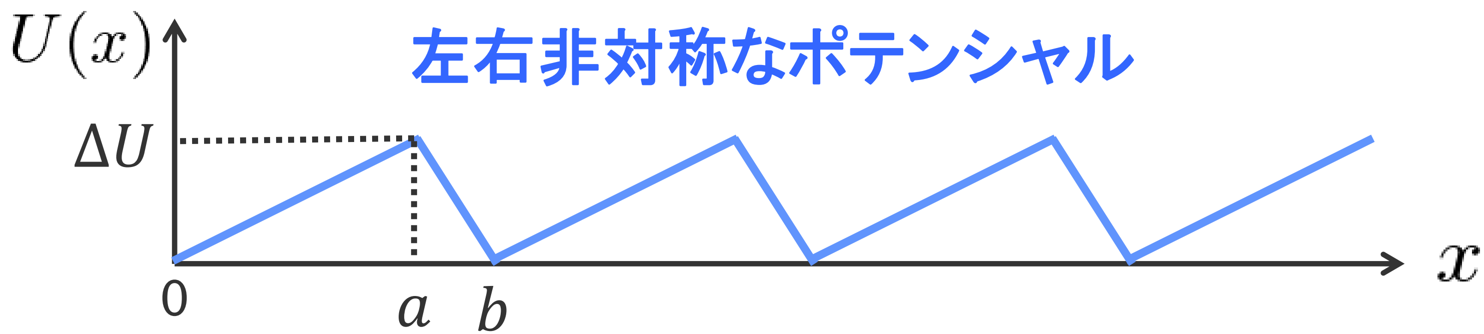
確率的爪車(ratchet)モデル

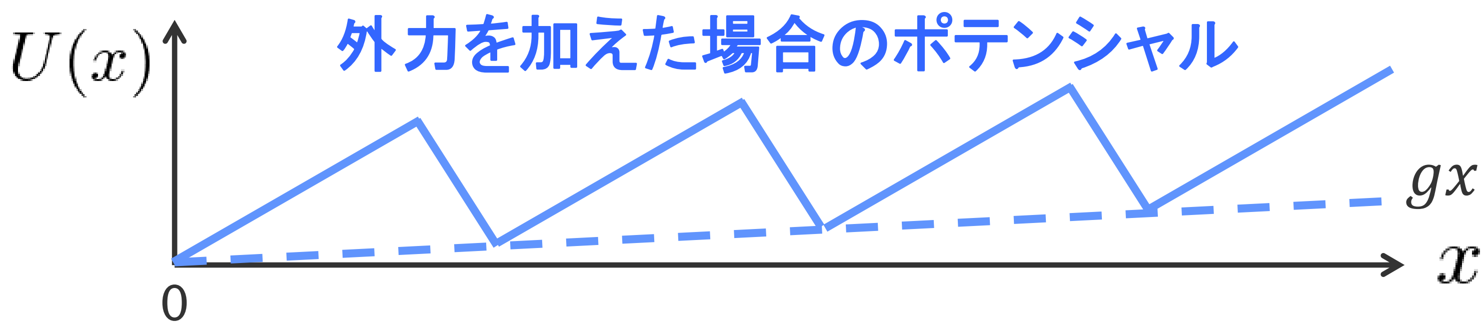
- ポテンシャル中のブラウン運動

$$\frac{\mathrm{d}x}{\mathrm{d}t} = -\frac{\mathrm{d}}{\mathrm{d}x}U(x) + z$$

$$\tau\frac{\mathrm{d}z}{\mathrm{d}t} = -z + \xi$$

$$\Longrightarrow \langle z(t)z(0)\rangle = \frac{D}{\tau}e^{-|t|/\tau}$$





Ratchet モデルを数値的に解く

- 数値スキーム

$$x_{n+1} = x_n + \Delta t[f(x_n) + z_n]$$

$$z_{n+1} = z_n - (\Delta t z_n - \sqrt{2D\Delta t}\mathcal{N}(0,1))/\tau$$

$$f(x) = \begin{cases} -\Delta U/a - g & (\mathrm{mod}(x,b) < a) \\ \Delta U/(b-a) - g & (\text{otherwise}) \end{cases}$$

入力例

	A	B	C	D
1	Brownian ratchet			
	Parameters			
	D =	1		
	tau =	10		
5	dU =	1		
	a =	9		
	b =	10		
	g =	0.05		
	f0 =	-0.16111		
10	f1 =	0.95		
	dt =	1		
	ddt =	1.414214		
	t	ξ	z	x
	0	0	0	0
15	1	1.7009	0.17009	-0.16111
	2	0.498987	0.20298	0.958979
	3	3.444318	0.527114	1.000847

(f0)=-B5/B6-B8

(f1)=B5/(B7-B6)-B8

(ddt)=SQRT(2*B3*B11)

(ξ) =$B$12*NORM.S.INV(RAND())

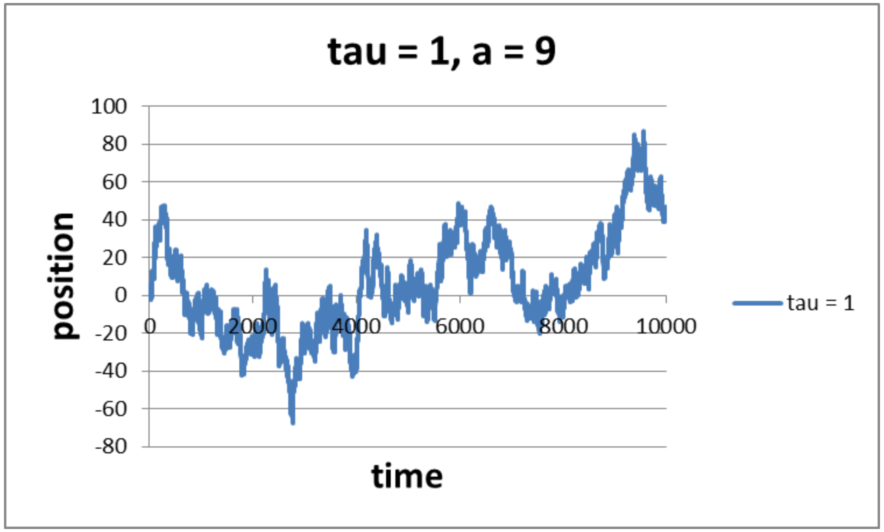
(z) =C14-($B$11*C14-B15)/$B$4

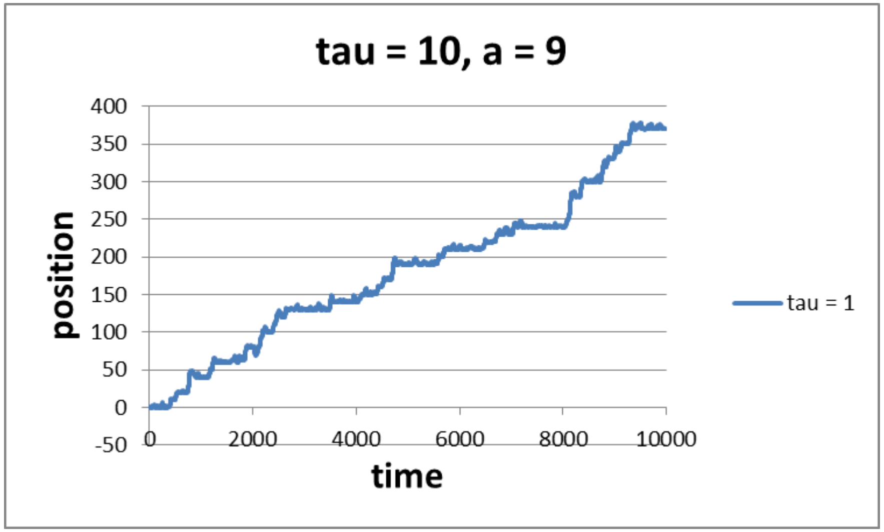
(x) =D14+$B$11*(
IF(MOD(D14,$B$7)
<$B$6,$B$9,$B$10)
+C14)

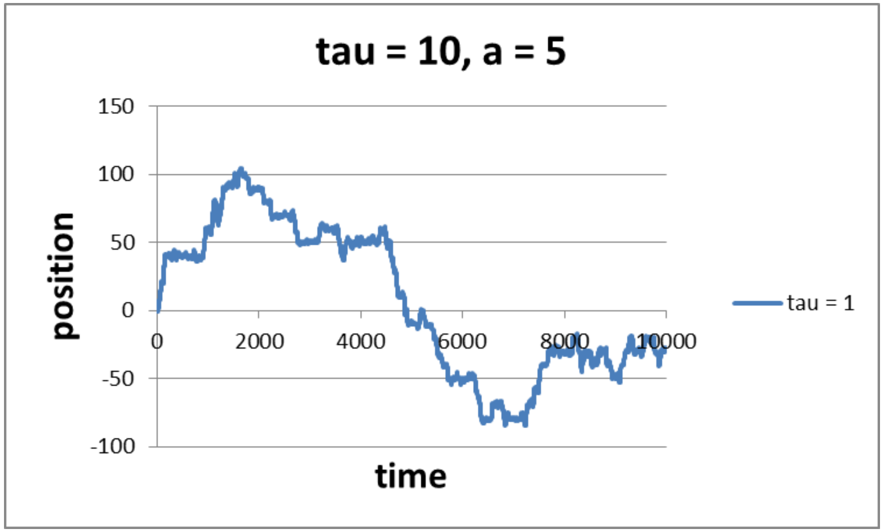
演習課題

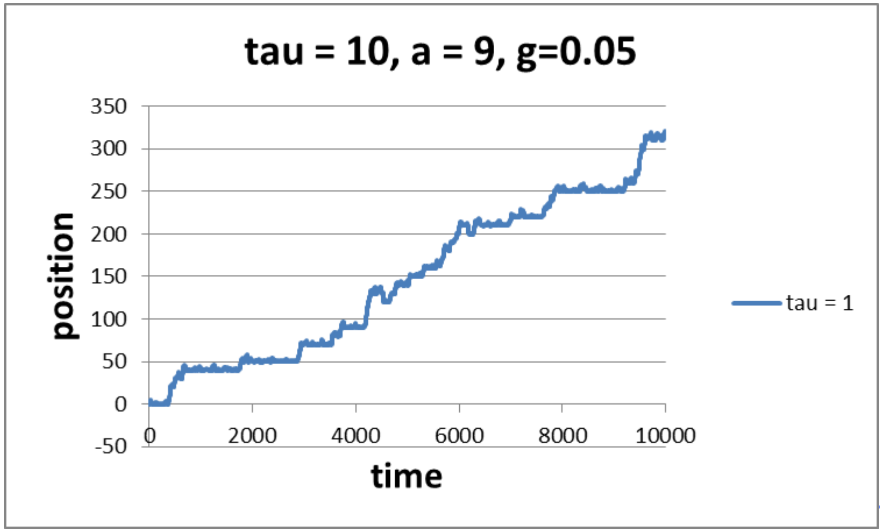
- Ratchet モデルの数値シミュレーションを行う。
  - 外力のない系（$g = 0$）で（平均的に）一方向に進む/進まないパラメータを見つける。
  - 外力のある系（$g \neq 0$）で同様のことを行う。
- (発展）パラメータ（例えば g, τ）を変えて、(x, t)のグラフが右上がり（＋1）、ほぼ平ら（0）、右下がり（－1）になるか調べ（それぞれのパラメータに対しF9キーにより何度か再計算した結果から判断する）、(g, τ)平面上に記す。

結果(例)

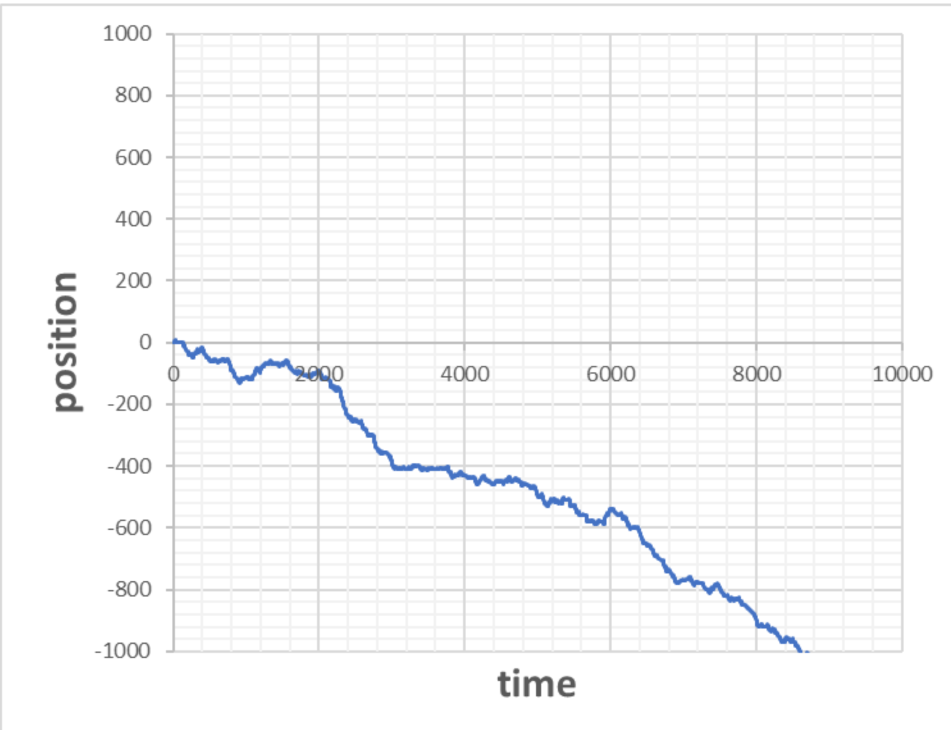

tau = 1, a = 9
position
time
tau = 1


tau = 10, a = 9
position
time
tau = 1


tau = 10, a = 5
position
time
tau = 1


tau = 10, a = 9, g=0.05
position
time
tau = 1

結果(発展)



(g, tau)				
10	1	1	1	0
9	1	1	1	0
8	1	1	0	-1
7	1	1	0	-1
6	1	0	-1	-1
5	1	0	-1	-1
4	0	0	-1	-1
3	0	-1	-1	-1
2	0	-1	-1	-1
1	0	-1	-1	-1
	0	0.05	0.1	0.15

(D=1, dU=1, a=9, b=10, dt=1)

参考文献

- 参考文献:東京大学教養学部統計学教室編「統計学入門」(東京大学出版会, 1991)
- 太田隆夫「非平衡系の物理学」(裳華房、2000)
- 江沢　洋　「だれが原子をみたか」岩波現代文庫(岩波書店、2013)
- 米沢富美子「ブラウン運動」物理学One Point 27(共立出版、1986)