Data Driven Models for Engineering Problems
Exercise 0

Konstantinos Papadakis
DSML 03400149
k.i.papadakis@gmail.com

March 2022

1 Exercises

Let $X(t_n)$ be a random walk where each step is either 1 or -1 with equal probability.

Exercise 1.1. $\mathrm{E}[X(t_n)] = 0$

Proof. Let $\xi(t_1), \dots, \xi(t_n)$ be the steps up to t_n. By assumption, these are independent and take values in $\{-1, 1\}$ with equal probability $\frac{1}{2}$.

$$\mathrm{E}[\xi_n] = (-1) \cdot P(\xi_n = -1) + 1 \cdot P(\xi_n = 1) = (-1) \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = 0$$

We have $X(t_n) = \xi(t_1) + \cdots + \xi(t_n)$, therefore

$$\mathrm{E}[X(t_n)] = \mathrm{E}[\xi(t_1)] + \cdots + \mathrm{E}[\xi(t_n)] = 0 + \cdots + 0 = 0$$

□

Exercise 1.2. $\mathrm{V}[X(t_n)] = n$

Proof.

$$\begin{aligned}
\mathrm{V}[\xi(t_n)] &= \mathrm{E}[\xi(t_n)^2] - \mathrm{E}[\xi(t_n)]^2 = \mathrm{E}[\xi(t_n)^2] \\
&= (-1)^2 \cdot P(\xi_n = -1) + 1^2 \cdot P(\xi_n = 1) \\
&= 1 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = 1
\end{aligned}$$

Because $\xi(t_n)$ are independent, we have that

$$\mathrm{V}[X(t_n)] = \mathrm{V}[\xi(t_1)] + \cdots + \mathrm{V}[\xi(t_n)] = 1 + \cdots + 1 = n$$

□

Exercise 1.3. *The process is not ergodic.*

Proof. Let $\hat{\mu}_X := \frac{1}{T}\sum_{i=0}^{T} X(t_n)$ be the time average estimate. We will show that this random variable does not converge to the (constant) ensemble average $\mu_x = 0$.

$$\mathrm{E}[\hat{\mu}_x] = \mathrm{E}\left[\frac{1}{T}\sum_{i=0}^{T} X(t_n)\right] = \frac{1}{T}\sum_{i=0}^{T} \mathrm{E}[X(t_n)] = 0$$

This shows that the time average estimate's mean is equal to μ_x, but as we shall prove, it does not converge to a *constant*.

$$\begin{aligned}
\mathrm{V}[\mu_X] &= \mathrm{E}[\mu_x^2] - \mathrm{E}[\mu_x]^2 = \mathrm{E}[\mu_x^2] \\
&= \frac{1}{T^2}\,\mathrm{E}\left[\left(\sum_{i=1}^{T} X(t_i)\right)^2\right] \\
&= \frac{1}{T^2}\,\mathrm{E}\left[\sum_{i=1}^{T}\sum_{j=1}^{T} X(t_i)X(t_j)\right] \\
&= \frac{1}{T^2}\sum_{i=1}^{T}\sum_{j=1}^{T} \mathrm{E}[X(t_i)X(t_j)] \\
&= \frac{1}{T^2}\sum_{i=1}^{T}\sum_{j=1}^{T} \mathrm{Cov}[X(t_i), X(t_j)] + \mathrm{E}[X(t_i)]\,\mathrm{E}[X(t_j)] \\
&= \frac{1}{T^2}\sum_{i=1}^{T}\sum_{j=1}^{T} \mathrm{Cov}[X(t_i), X(t_j)]
\end{aligned}$$

Since $\xi(t_i)$ are independent and with variance 1, we have that

$$\begin{aligned}
\mathrm{Cov}[X(t_i), X(t_j)] &= \mathrm{Cov}[\sum_{r=1}^{i} \xi(t_r), \sum_{s=1}^{j} X(t_s)] \\
&= \sum_{r=s} \mathrm{Cov}[\xi(t_r), \xi(t_s)] + \sum_{r\neq s} \mathrm{Cov}[\xi(t_r), \xi(t_s)] \\
&= \sum_{k=1}^{\min\{i,j\}} \mathrm{Cov}[\xi(t_k), \xi(t_k)] \\
&= \sum_{k=1}^{\min\{i,j\}} 1 \\
&= \min\{i, j\}
\end{aligned}$$

Therefore,

$$V[\mu_X] = \frac{1}{T^2} \sum_{i=1}^{T} \sum_{j=1}^{T} \min\{i, j\}$$

We now show that $\sum_{i=1}^{T} \sum_{j=1}^{T} \min\{i, j\} = \sum_{k=1}^{T} k^2$. Evaluate the summands of the left hand side in the following table:

$$\begin{array}{cccccc}
1 & 1 & 1 & 1 & \dots & 1 \\
1 & 2 & 2 & 2 & \dots & 2 \\
1 & 2 & 3 & 3 & \dots & 3 \\
1 & 2 & 3 & 4 & \dots & 4 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
1 & 2 & 3 & 4 & \dots & T
\end{array}$$

We can view this table as stacked 3-dimensional unit cubes. The number at position (i, j) represents the number of cubes that are stacked vertically at this position. At height 1 we have T^2 cubes, at height 2 we have $(T-1)^2$ cubes, $\dots$, at height T we have 1^2 cubes. This is proves the equality visually.

We now have

$$\begin{aligned}
V[\mu_X] &= \frac{1}{T^2} \sum_{i=1}^{T} \sum_{j=1}^{T} \min\{i, j\} \\
&= \frac{1}{T^2} \sum_{k=1}^{T} k^2 \\
&= \frac{1}{T^2} \frac{T(T+1)(2T+1)}{6} \\
&= \frac{1}{6}(1 + \frac{1}{T})(2T+1) \xrightarrow{T \to \infty} \infty
\end{aligned}$$

where we used the well-known formula for the sum of squares of the first n natural numbers.

This shows that the variance of μ_X does not become 0, and thus μ_X doesn't converge to a constant. $\square$

2 Code

We can also write Matlab code that attempts to estimate the values proved above.

```
function [means, vars, mean_muX, var_muX] = exercise_0(n_steps,
    n_samples)

    n_steps = n_steps + 1;  % Including the initial position (0-th
    step)
```

```
    % X(t_n) of each sample.
    pos = zeros(n_samples, 1);  % Initial positions
    % E[X(t_n)] of each time step n. True value is 0 for all n.
    means = zeros(n_steps, 1);
    means(1) = mean(pos);
    % V[X(t_n)] of each time step n. True value is n.
    vars = zeros(n_steps, 1);
    vars(1) = var(pos);
    % Average over all time (for each sample). Random variable muX.
    muX = pos;

    for i = 2:n_steps
        step = 2 * (rand(n_samples, 1) > 0.5) - 1;  % random +-1s
        pos = pos + step;
        means(i) = mean(pos);
        vars(i) = var(pos);
        muX = muX + (pos - muX) / i;  % running update
    end
    % Expected value of average over all time. True value is 0.
    mean_muX = mean(muX);
    % Variance of average over all time.
    % True value is 1/6 * (n_steps + 1) / n_steps * (2*n_steps + 1)
    var_muX = var(muX);
end
```

Listing 1: Matlab simulation