
Consider the following Markov Process on the integers $Z = \{i : -\infty < i < \infty\}$. The process $x_n$ starts from 0 at time $n = 0$ and at each step moves one unit to the right with probability $p$ or one unit to the left with probability $q = 1 - p$. The choices at successive steps are made independently, but always with the same probabilities $p$ and $q$ for moving right or left. This clearly defines a Markov Process on $Z$ with

$$\pi_{i,i+1} = p; \quad \pi_{i,i-1} = q; \quad \text{and} \quad \pi_{i,j} = 0 \quad \text{for} \quad |i - j| \neq 1$$

the $n$ step transition probabilities are easy to calculate.

$$\pi_{i,j}^{(n)} = 0 \quad \text{unless} \quad |j - i| \leq n \quad \text{and} \quad j - i = n \quad (\text{mod} \ 2)$$

$$\pi_{i,j}^{(n)} = \binom{n}{r} p^r q^{n-r} \quad \text{where} \quad 2r - n = j - i$$

One can obtain this formula by noting that out of the $n$ steps $r$ were to the right and the remaining $n - r$ were to the left $x_n = i + r - (n - r) = i + 2r - n$ and $x_n = j$ if $j - i = 2r - n$. This is possible only if $j - i = n$ (mod 2), and the probability is given by the binomial distribution. The law of large numbers for the binomial says that for large $n$ the probability is nearly 1 that $\frac{r}{n}$ is close to $p$. More precisely, for any $\epsilon > 0$,

$$\lim_{n \to \infty} \sum_{r : \left| \frac{r}{n} - p \right| \leq \epsilon} \binom{n}{r} p^r q^{n-r} = 1$$

This means that for any bounded continuous function $f$ on $R$

$$\lim_{n \to \infty} E[f\left(\frac{x_n}{n}\right)|x_0 = 0] = f(p - q)$$

If we denote by $P$ the operator

$$(Pu)(i) = pu(i + 1) + (1 - p)u(i - 1)$$

Our statement concerns the function $u_n(i) = f\left(\frac{i}{n}\right)$ and the claim is

$$\lim_{n \to \infty} (P^n u_n)(0) = p - q$$

Let us see if we can figure out why this is true. Let us rescale space and time by a step size of $h = \frac{1}{n}$. Then

$$(Pu)(x) = pu(x + h) + qu(x - h)$$
Let us define \( u(kh, x) = (P^k u)(x) \). We see that
\[
u((k+1)h, x) = pu(kh, x + h) + qu(kh, x - h)
\]
Or
\[
\frac{1}{h}[u((k+1)h, x) - u(kh, x)] = \frac{1}{h}p[u(kh, x + h) - u(kh, x)] + \frac{1}{h}q[u(kh, x - h) - u(kh, x)]
\]
Passing to the limit as \( h \to 0 \) we get
\[
u_t = (p - q)u_x
\]
The solution with \( u(0, x) = f(x) \) is given by
\[
u(t, x) = f(x + (p - q)t)
\]
We are interested in \( u(1, 0) \) which is \( f(p - q) \). The law of large numbers for the binomial is just the approximation of a first order PDE by difference equations.

Let us now consider the case where \( p = q = \frac{1}{2} \). Now \( P^n u_n(x) \to f(x) \), so nothing much happens. The space scale has to be \( \frac{1}{\sqrt{n}} \) to get something significant. In the probabilistic setting we are looking at \( \frac{x_n}{\sqrt{n}} \) which satisfies the central limit theorem and the correct behavior is with \( u(x) = f(\frac{x}{\sqrt{n}}) \)
\[
(P^n u_n)(0) \to \int \frac{1}{\sqrt{2\pi}} f(y)e^{-\frac{y^2}{2}} dy
\]
This is just as before except we get
\[
\frac{1}{h}[u((k+1)h, x) - u(kh, x)] = \frac{1}{2h}[u(kh, x + \sqrt{h}) - u(kh, x)] + \frac{1}{2h}q[u(kh, x - \sqrt{h}) - u(kh, x)]
\]
and passing to the limit as \( h \to 0 \), we get the heat equation
\[
u_t = \frac{1}{2}u_{xx}
\]
with \( u(0, x) = f(x) \). The solution is given by
\[
u(t, x) = \int \frac{1}{\sqrt{2\pi t}} f(y)e^{-\frac{(y-x)^2}{2t}} dy
\]
and \( u(1, 0) \) is then
\[
\int \frac{1}{\sqrt{2\pi}} f(y)e^{-\frac{y^2}{2}} dy
\]
In other words, the central limit theorem for the binomial (with \(p = q = \frac{1}{2}\)) can be interpreted as the convergence of the solution of

\[
u(t + h, x) = \frac{1}{2}[u(t, x + \sqrt{h}) + u(t, x - \sqrt{h})]
\]

to the corresponding solution of the heat equation (1).

We can have time varying continuously, and consider a Markov process on \(Z\) with transition rates

\[
a_{i,i+1} = a_{i,i-1} = \frac{1}{2}; \quad a_{i,i} = -1; \quad \text{and} \quad a_{i,j} = 0 \quad \text{otherwise}.
\]

The expectation

\[
u(t, i) = E[f(x(t)|x(0) = i]
\]

will satisfy

\[
\frac{du(t, i)}{dt} = \frac{1}{2}[u(t, i + 1) + u(t, i - 1) - 2u(t, i)]; \quad u(0, i) = f(i)
\]

If we rescale space by \(\sqrt{h}\) and time by \(h\), the equations become

\[
\frac{du(t, x)}{dt} = \frac{1}{2h}[u(t, x + \sqrt{h}) + u(t, x - \sqrt{h}) - 2u(t, x)]; \quad u(0, x) = f(x)
\]

which again converges to the solution of the heat equation. What we have is again a central limit theorem for the distribution of \(\frac{x(t)}{\sqrt{t}}\) as \(t \to \infty\).

**Examples.**

1. We will show that the distribution of \(x(t)\) is the distribution of the difference \(X_1 - X_2\) of two independent Poisson random variables with parameters \(\frac{t}{2}\). Let us try \(f(i) = e^{\lambda i}\). The solution \(u(t, i)\) can be obtained by separation of variables. Set \(u(t, i) = e^{\lambda i}g(\lambda, t)\). Then,

\[
\frac{dg}{dt} = \frac{1}{2}[e^\lambda + e^{-\lambda} - 2]g; \quad g(\lambda, 0) = 1
\]

will do it. This gives

\[
g(\lambda, t) = \exp[t\frac{1}{2}((e^\lambda - 1) + (e^{-\lambda} - 1))] = E[e^{\lambda(X_1 - X_2)}]
\]

with \(X_1, X_2\) having independent Poisson distributions with parameter \(\frac{t}{2}\).
Suppose $\pi_h(x, y)dy$ is the transition density of a Markov chain, with $h$ representing the time step. Let us make the following assumptions:

$$\sup_x |\int (y - x)\pi_h(x, y)dy| = o(h)$$

$$\sup_x |\int (y - x)^2\pi_h(x, y)dy - h| = o(h)$$

$$\sup_x \int (y - x)^4\pi_h(x, y)dy = o(h)$$

Then the function $u_h(n, x) = E[f(x_n)|x_0 = x]$ converges to the solution $u$ of the heat equation (1) provided $nh \to t$.

**Proof:** Let us start with the solution $u(t, x)$ of the heat equation which we will assume is a smooth function of $t$ and $x$. Let

$$\Delta_k = E[u(t - (k + 1)h, x_{k+1})|x_0 = x] - E[u(t - kh, x_k)|x_0 = x]$$

Assuming $nh = t$, this is a telescoping sum and

$$\sum_{k=0}^{n-1} \Delta_k = u_h(n, x) - u(t, x)$$

It is therefore sufficient to prove that each

$$\Delta_k = o(h)$$

If we expand by Taylor’s formula

$$u(t - (k + 1)h, y) - u(t - kh, x) = -hu_t(t - kh, x) + u_x(t - kh, x)(y - x)$$

$$+ \frac{1}{2} u_{xx}(t - kh, x)(y - x)^2 + \text{Remainder}$$

We can estimate

$$E[u(t - (k + 1)h, x_{k+1}) - u(t - kh, x_k)|x_k = x]$$

by $o(h)$, because upon integrating with $\pi_h(x, y)$, with errors that are $o(h)$, we get $h(u_t - \frac{1}{2} u_{xx}) = 0$. The remainder term can be estimated by

$$\int |y - x|^3\pi_h(x, y)dy \leq \left( \int |y - x|^2\pi_h(x, y)dy \right)^{\frac{1}{2}} \left( \int |y - x|^4\pi_h(x, y)dy \right)^{\frac{1}{2}}$$

$$\leq o(h)$$

To go from $E[Q|x_k]$ to $E[Q|x_0]$ is easy because the conditional expectation of anything that is $o(h)$ is still $o(h)$. 

4
Remarks:

1. We have assumed that \( \pi_h(x, y)dy \) is given by a density only for convenience. In principle \( \pi_h(x, \cdot) \) is just the probability distribution of \( x_1 \) given \( x(0) = x \), and does not have to be given by a density. It can be a discrete distribution as well.

2. The first two assumptions say that to within \( o(h) \) the infinitesimal ‘mean’ is 0 and the infinitesimal ‘variance’ is \( h \).

3. The third condition is important and needs to be understood. A random variable \( X \) with mean 0 and variance \( h \) can come in different shapes. For example \( X \) can be \( \pm \sqrt{h} \) with probability \( \frac{1}{2} \) each. Or \( X \) can be 0 with probability \( 1 - h \) and \( \pm 1 \) with probability \( \frac{h}{2} \) each. In both cases the mean and variance check out. However \( E[X^4] = h^2 \) in the first case and \( E[X^4] = h \) in the second. The clue is in the calculation of

\[
\lim_{h \to 0} \frac{1}{h} E[f(X) - f(0)] = \frac{1}{2} f_{xx}(0)
\]

in the first case and

\[
\lim_{h \to 0} \frac{1}{h} E[f(X) - f(0)] = \frac{1}{2} [f(1) + f(-1) - 2f(0)]
\]

in the second.

Examples:

2*. (Compare with Ex 3* of section 2). Suppose \( \pi_h(x, y) \) satisfies

\[
\sup_x | \int (y - x) \pi_h(x, y)dy - hb(x) | = o(h)
\]

and

\[
\sup_x | \int (y - x)^2 \pi_h(x, y)dy | = o(h)
\]

for some nice smooth bounded function \( b(x) \), prove the ‘law of large numbers’

\[
\lim_{h \to 0} P[|x_n - g(t)| \geq \epsilon | x_0 = x] \to 0
\]

where \( g(t) \) is the value of the solution at time \( t \) of

\[
\frac{dg}{ds} = b(g(s)); \ g(0) = x
\]

Note that it is sufficient to prove

\[
\lim_{h \to 0} \lim_{n \to \infty} \frac{u_h(n, x)}{nh} = \lim_{h \to 0} \lim_{n \to \infty} E[f(x_n) | x_0 = x] = u(t, x) = f(g(t))
\]

where \( u \) solves \( u_t = b(x)u_x \) with \( u(0, x) = f(x) \).