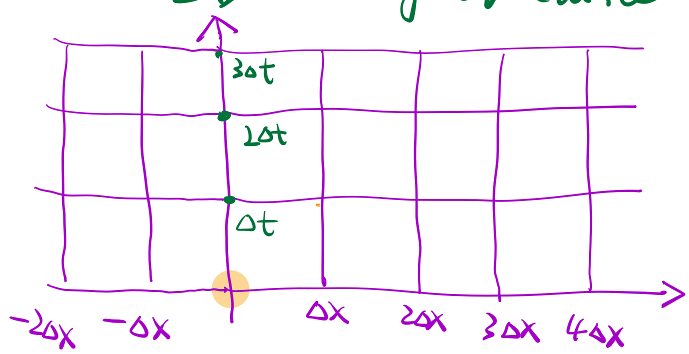
Concepts & Examples and more

- X & Y are random variables taking values in $\{x_1, x_2, \dots\}$ & $\{y_1, y_2, \dots\}$
- Joint Probability $P(X=x_i, Y=y_j)$
- Conditional Probability $P(X=x_i \mid Y=y_j) = \dfrac{P(X=x_i, Y=y_j)}{P(Y=y_j)}$

Example : Random Walk (Einstein 1905 paper)

2D rectangular lattice defined by $(m\Delta x, n\Delta t)$



1) A particle starts at $x=0$ at time $t=0$

2) At time step n, the position is X_n

$$X_{n+1} = \begin{cases} X_n - \Delta x, & \text{with probability } \frac{1}{2} \\ X_n + \Delta x, & \text{with prob. } \frac{1}{2} \end{cases}$$

3) So $X_0 = 0$ is deterministic

$\left.\begin{matrix} X_1 \\ X_2 \\ X_3 \\ \vdots \end{matrix}\right\}$ are random variables

Examples of Joint & conditional probability

$P(X_1 = 2\Delta x, X_2 = 3\Delta x) = 0$

$$\begin{cases} P(X_1 = -\Delta x, X_2 = -2\Delta x) = \frac{1}{4} \\ P(X_2 = -2\Delta x \mid X_1 = -\Delta x) = \frac{1}{2} \\ P(X_1 = -\Delta x) = \frac{1}{2} \end{cases}$$

consistent with $P(X=x \mid Y=y) = \dfrac{P(X=x, Y=y)}{P(Y=y)}$

- Expectation $E(X) = \sum_i x_i \cdot \text{Prob}(X = x_i)$
- Conditional Expectation

1) $E(X \mid Y = y_k)$ is a number

$$E(X \mid Y = y_k) = \sum_i x_i \cdot \text{Prob}(X = x_i \mid Y = y_k)$$

2) $E(X \mid Y)$ is a random variable depending on Y

$$E(X \mid Y) = \sum_i x_i \,\text{Prob}(X = x_i \mid Y)$$

$$\begin{aligned} E(X_2 \mid X_1 = \Delta x) &= 2\Delta x \cdot \text{Prob}(X_2 = 2\Delta x \mid X_1 = \Delta x) \\ &\quad + 0 \cdot \text{Prob}(X_2 = 0 \mid X_1 = \Delta x) \\ &= 2\Delta x \cdot \frac{1}{2} = \Delta x \end{aligned}$$

Similarly $E(X_n \mid X_{n-1} = m\Delta x) = m \cdot \Delta x$

$$E(X_2 \mid X_1) = \sum_{m=-\infty}^{+\infty} m \cdot \Delta x \cdot \text{Prob}(X_2 = m\Delta x \mid X_1)$$

is a function of X_1

thus a random variable

$$E(X_2 \mid X_1) = \begin{cases} \Delta x, & X_1 = \Delta x \text{ with prob. } \frac{1}{2} \\ -\Delta x, & X_1 = -\Delta x \text{ with prob. } \frac{1}{2} \end{cases}$$

$$E([X_2]^3 \mid X_1) = \begin{cases} (2\Delta x)^3 \frac{1}{2} + 0 \cdot \frac{1}{2} & \text{if } X_1 = \Delta x \\ (-2\Delta x)^3 \cdot \frac{1}{2} + 0 \cdot \frac{1}{2} & \text{if } X_1 = -\Delta x \end{cases}$$

$$= \begin{cases} 4\Delta x^3, & \text{with prob. } \frac{1}{2} \\ -4\Delta x^3, & \text{with prob. } \frac{1}{2} \end{cases}$$

$$E(X_3 \mid X_2, X_1)$$

$$= \sum_{m=-\infty}^{\infty} m\Delta x \cdot \text{Prob}(X_3 = m\Delta x \mid X_2, X_1)$$

$$\begin{aligned} = \ & 3\Delta x \cdot \text{Prob}(X_3 = 2\Delta x \mid X_2, X_1) \\ & + 2\Delta x \cdot \text{Prob}(X_3 = 2\Delta x \mid X_2, X_1) \\ & + \Delta x \cdot \text{Prob}(X_3 = \Delta x \mid X_2, X_1) \\ & + 0 \cdot \text{Prob}(X_3 = 0 \mid X_2, X_1) \\ & - \Delta x \cdot \text{Prob}(X_3 = -\Delta x \mid X_2, X_1) \\ & - 2\Delta x \cdot \text{Prob}(X_3 = -2\Delta x \mid X_2, X_1) \quad \text{is a random variable} \\ & - 3\Delta x \cdot \text{Prob}(X_3 = -3\Delta x \mid X_2, X_1) \quad \text{depending on } X_2, X_1 \end{aligned}$$

- Martingale is a sequence $\{X_k\}_{k=1}^{\infty}$ satisfying

① X_k is a R.V. (random variable)

② $E(|X_k|) < +\infty$, $\forall k$

③ $E(X_{k+1} \mid X_k, \underbrace{X_{k-1}, \cdots, X_1}_{\text{forgetting history}}\} = X_k$

Claim $E(X_3 \mid X_2, X_1) = X_2$

$$
\begin{aligned}
& E(X_3 \mid X_2, X_1) \\
= {} & 3\Delta x \cdot \text{Prob}(X_3 = 2\Delta x \mid X_2, X_1) \\
& + 2\Delta x \cdot \text{Prob}(X_3 = 2\Delta x \mid X_2, X_1) \\
& + \Delta x \cdot \text{Prob}(X_3 = \Delta x \mid X_2, X_1) \\
& + 0 \cdot \text{Prob}(X_3 = 0 \mid X_2, X_1) \\
& - \Delta x \cdot \text{Prob}(X_3 = -\Delta x \mid X_2, X_1) \\
& - 2\Delta x \cdot \text{Prob}(X_3 = -2\Delta x \mid X_2, X_1) \\
& - 3\Delta x \cdot \text{Prob}(X_3 = -3\Delta x \mid X_2, X_1)
\end{aligned}
$$

$$
= \begin{cases}
3\Delta x \cdot \frac{1}{2} + \Delta x \frac{1}{2} = 2\Delta x & \text{if } X_2 = 2\Delta x \\
2\Delta x \frac{1}{2} + 0 \cdot \frac{1}{2} = \Delta x & \text{if } X_2 = \Delta x \\
\Delta x \frac{1}{2} + (-\Delta x) \frac{1}{2} = 0 & \text{if } X_2 = 0 \\
0 \cdot \frac{1}{2} + 2\Delta x \frac{1}{2} = \Delta x & \text{if } X_2 = -\Delta x \\
-3\Delta x \cdot \frac{1}{2} + (-\Delta x) \frac{1}{2} = -2\Delta x & \text{if } X_2 = -2\Delta x
\end{cases}
$$

is the same as X_2 !

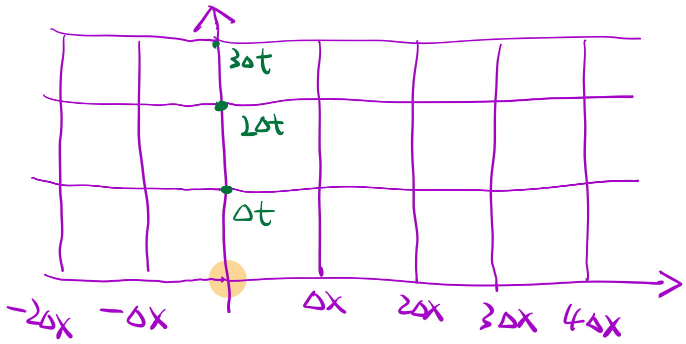
Similarly, $E(X_{n+1} \mid X_n, X_{n-1}, \cdots, X_1) = X_n$

So this is a martingale because

$$E(|X_k|) \le k\Delta x < +\infty$$

This example is related to many things like heat equation, centered difference, Gaussian distribution, Brownian motion, etc.

More on this example



Let $p(m, n)$ denote the probability that

the particle is at position $m\Delta x$ at time $n\Delta t$. Then

$$p(m, 0) = \begin{cases} 0 & m \neq 0 \\ 1 & m = 0. \end{cases}$$

Also

$$p(m, n+1) = \frac{1}{2}p(m-1, n) + \frac{1}{2}p(m+1, n),$$

and hence

$$p(m, n+1) - p(m, n) = \frac{1}{2}(p(m+1, n) - 2p(m, n) + p(m-1, n)).$$

Now assume

$$\frac{(\Delta x)^2}{\Delta t} = D \quad \text{for some positive constant } D.$$

This implies

$$\frac{p(m, n+1) - p(m, n)}{\Delta t} = \frac{D}{2}\left(\frac{p(m+1, n) - 2p(m, n) + p(m-1, n)}{(\Delta x)^2}\right).$$

Let $\Delta t \to 0$, $\Delta x \to 0$, $m\Delta x \to x$, $n\Delta t \to t$, with $\frac{(\Delta x)^2}{\Delta t} \equiv D$. Then presumably $p(m, n) \to f(x, t)$, which we now interpret as the probability density that particle

is at x at time t. The above difference equation becomes formally in the limit

$$f_t = \frac{D}{2} f_{xx},$$

and so we arrive at the diffusion equation again.

- Variance of a R.V.

expectation $E[\cdot]$ is linear

$$V(X) = E[(X - E(X))^2] = E[X^2 - 2E(X)X + E(X)^2]$$

$$= E[X^2] - E[2E(X)\cdot X] + E(X)^2 = E[X^2] - E[X]^2$$

MATHEMATICAL JUSTIFICATION. A more careful study of this technique of passing to limits with random walks on a lattice depends upon the Laplace–De Moivre Theorem.

As above we assume the particle moves to the left or right a distance Δx with probability $1/2$. Let $X(t)$ denote the position of particle at time $t = n\Delta t \quad (n = 0, \dots)$. Define

$$S_n := \sum_{i=1}^{n} X_i,$$

where the X_i are independent random variables such that

$$\begin{cases} P(X_i = 0) = 1/2 \\ P(X_i = 1) = 1/2 \end{cases}$$

for $i = 1, \dots$. Then $V(X_i) = \frac{1}{4}$.

Now S_n is the number of moves to the right by time $t = n\Delta t$. Consequently

$$X(t) = S_n \Delta x + (n - S_n)(-\Delta x) = (2S_n - n)\Delta x.$$

Note also

$$\begin{aligned} V(X(t)) &= (\Delta x)^2 V(2S_n - n) \\ &= (\Delta x)^2 4V(S_n) = (\Delta x)^2 4nV(X_1) \\ &= (\Delta x)^2 n = \frac{(\Delta x)^2}{\Delta t} t. \end{aligned}$$

Again assume $\frac{(\Delta x)^2}{\Delta t} = D$. Then

$$X(t) = (2S_n - n)\Delta x = \left(\frac{S_n - \frac{n}{2}}{\sqrt{\frac{n}{4}}} \right) \sqrt{n}\Delta x = \left(\frac{S_n - \frac{n}{2}}{\sqrt{\frac{n}{4}}} \right) \sqrt{tD}.$$

The Laplace–De Moivre Theorem thus implies

$$\begin{aligned} \lim_{\substack{n\to\infty \\ t = n\Delta t, \frac{(\Delta x)^2}{\Delta t} = D}} P(a \le X(t) \le b) &= \lim_{n\to\infty} \left(\frac{a}{\sqrt{tD}} \le \frac{S_n - \frac{n}{2}}{\sqrt{\frac{n}{4}}} \le \frac{b}{\sqrt{tD}} \right) \\ &= \frac{1}{\sqrt{2\pi}} \int_{\frac{a}{\sqrt{tD}}}^{\frac{b}{\sqrt{tD}}} e^{-\frac{x^2}{2}}\, dx \\ &= \frac{1}{\sqrt{2\pi Dt}} \int_a^b e^{-\frac{x^2}{2Dt}}\, dx. \end{aligned}$$

Once again, and rigorously this time, we obtain the $N(0, Dt)$ distribution. $\square$

Laplace–De Moivre

is a special case of central limit Thm

LEMMA. *Suppose the real–valued random variables $X_1, \dots, X_n, \dots$ are independent and identically distributed, with*

$$\begin{cases} P(X_i = 1) = p \\ P(X_i = 0) = q \end{cases}$$

for $p, q \geq 0$, $p + q = 1$. Then

$$E(X_1 + \cdots + X_n) = np$$
$$V(X_1 + \cdots + X_n) = npq.$$

THEOREM (Laplace–De Moivre). *Let $X_1, \dots, X_n$ be the independent, identically distributed, real–valued random variables in the preceding Lemma. Define the sums*

$$S_n := X_1 + \cdots + X_n.$$

Then for all $-\infty < a < b < +\infty$,

$$\lim_{n \to \infty} P\left(a \leq \frac{S_n - np}{\sqrt{npq}} \leq b\right) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}}\, dx.$$

A quick numerical demonstration for

$$p(m, n+1) - p(m, n) = \frac{1}{2}(p(m+1, n) - 2p(m, n) + p(m-1, n)).$$

Now assume

$$\frac{(\Delta x)^2}{\Delta t} = D \qquad \text{for some positive constant } D.$$

This implies

$$\frac{p(m, n+1) - p(m, n)}{\Delta t} = \frac{D}{2}\left(\frac{p(m+1, n) - 2p(m, n) + p(m-1, n)}{(\Delta x)^2}\right)$$

↳ This is Forward Euler + Centered Difference for $P_t = P_{xx}$ if $D = 2$

$$P_j^{n+1} = P_j^n - \frac{\Delta t}{\Delta x^2}\left[-P_{j+1}^n + 2P_j^n - P_{j-1}^n\right]$$

1) As a PDE scheme, need $\frac{\Delta t}{\Delta x^2} \leq \frac{1}{2}$ stability

2) As a gradient descent for $\min_{\vec{p}} f(\vec{p})$

$$f(\vec{p}) = \frac{1}{2} \vec{p}^T K \vec{p}$$

$$K = \frac{1}{\Delta x^2} \begin{pmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 2 \end{pmatrix}$$

$$\vec{p}_{n+1} = \vec{p}_n - \eta \nabla \vec{f}(\vec{p}_n)$$

$$= \vec{p}_n - \eta \cdot \frac{1}{\Delta x^2} K \vec{p}_n$$

$\eta < \frac{2}{L} \Rightarrow \frac{\eta}{\Delta x^2} < \frac{1}{2}$ because $L = \|K\| < 4$