Lecture-29

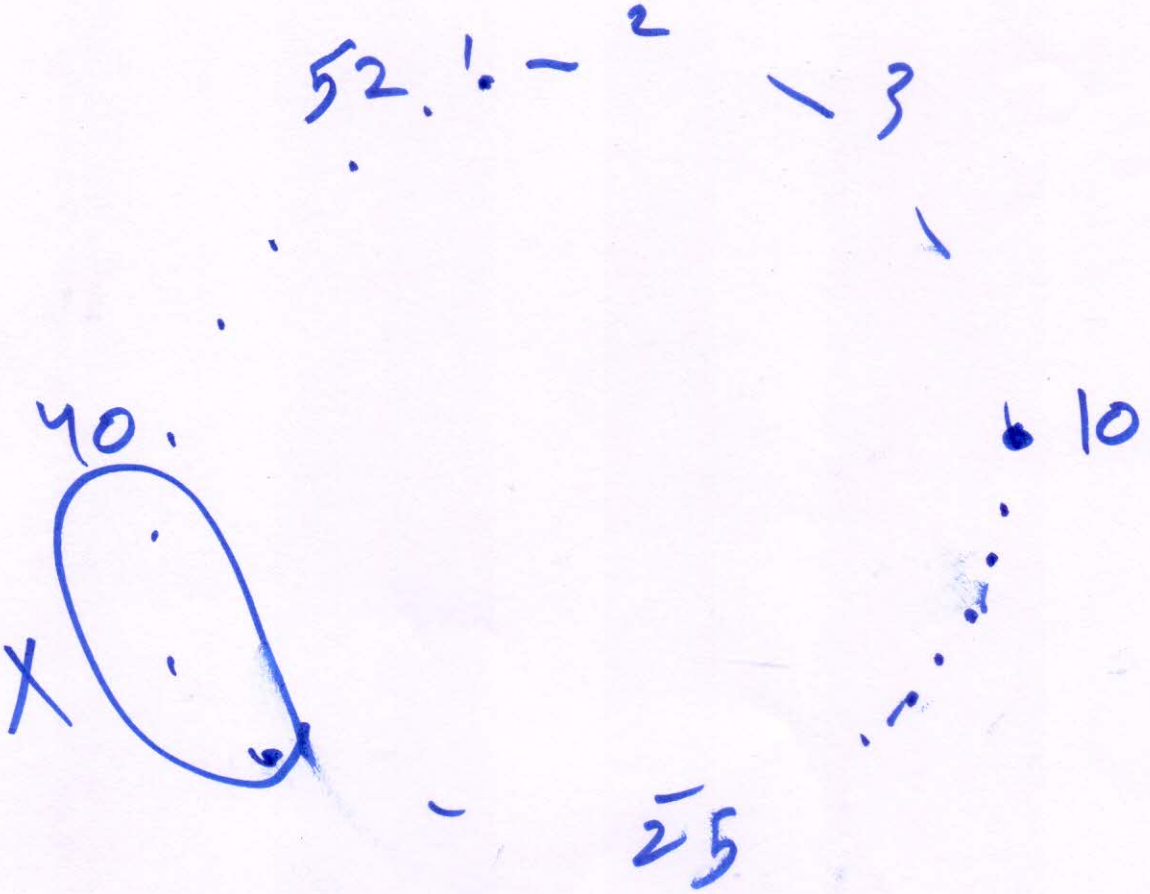
Recap
- → Random Walk
- → Quick Sort analysis

The Probabilistic Method

Applications in combinatorics & graph theory.

Prove existence.

E.g. 52 trees arranged in a circle. 15 Chipmunks live in these trees. Prove that there are 7 consecutive trees s.t. they together house at least 3 chipmunks.



Neighborhood of a tree = that tree and 6 more trees in the clockwise direction

$N(10) = \{10, 11, 12, 13, 14, 15, 16\}$

$N(50) = \{50, 51, 52, 1, 2, 3, 4\}$

Choose a random tree.

Let X be the no. of chipmunks in its neighborhood.

$E[X] > 2$. → Want to prove.

$$X_1, X_2, \vdots, X_{15} = \begin{cases} 1 & \text{if chipmunk no. } i \text{ is living in the n'hood of that random tree} \\ 0 & \text{o.w.} \end{cases}$$

$$\sum X_i = \underset{L}{X}$$

15 chipmunks

$$E[X] = E\left[\sum X_i\right]$$

$$= \sum_{i=1}^{15} E[X_i]$$

$$E[X_i] = 1 \cdot P(X_i = 1) + 0 \cdot P(X_i = 0)$$

$$= P(X_i = 1)$$

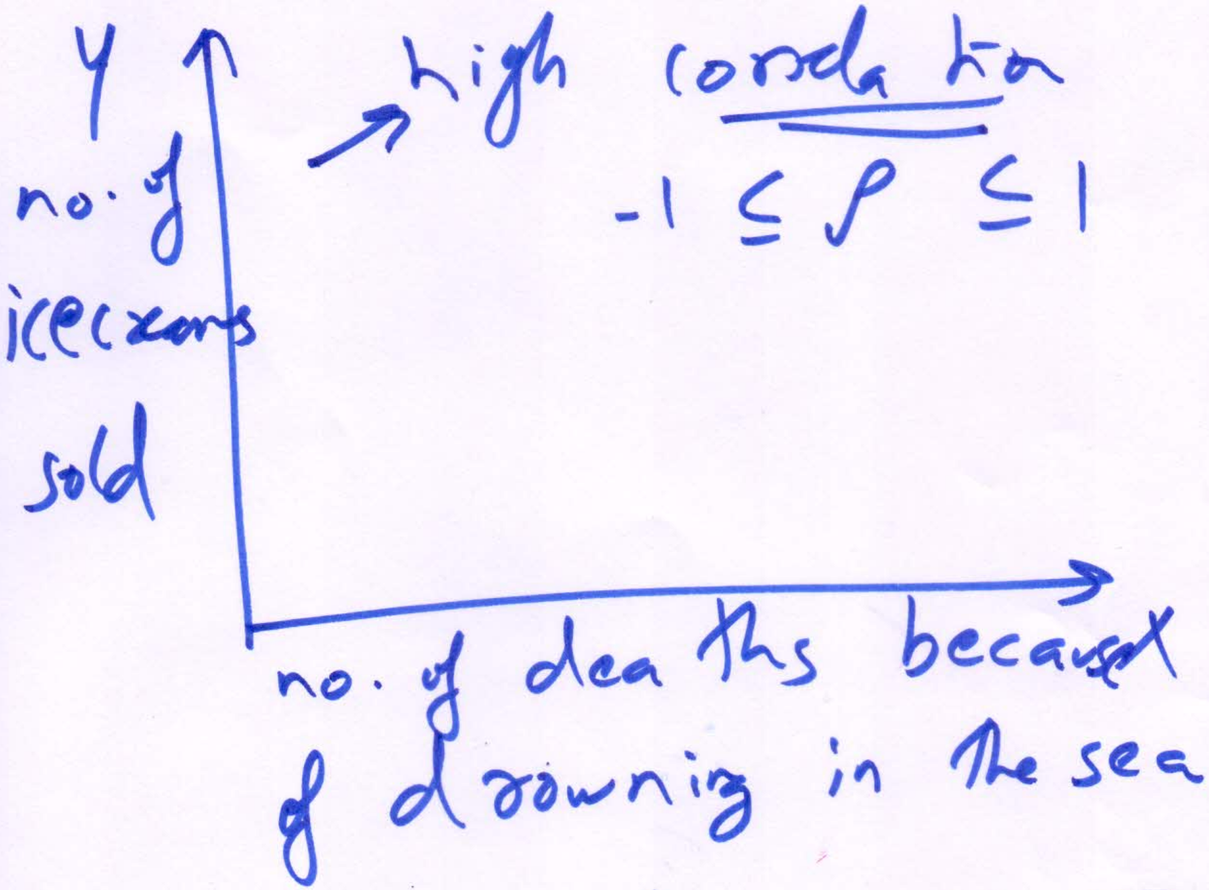
$$= \frac{7}{52}$$

$$E[X] = \sum_{i=1}^{15} \frac{7}{52} = \frac{15 \cdot 7}{52} = \frac{105}{52} > 2$$

Pioneered by Paul Erdös
The Probabilistic Method

Erdös number

Covariance & Correlation
Causality



$-1 \le \rho \le 1$

Correlation does NOT neccesarily imply causality.

Theorem:

X, Y: independent random variables.

Then

$$E[g(X)h(Y)] = E[g(X)]\,E[h(Y)]$$

$$\Downarrow$$

$$E[XY] = E[X]\,E[Y]$$

$$\iint g(x)\,h(y)\,\underset{f_X(x)\,f_Y(y)}{\underline{f(X,Y)}}\,dx\,dy \xrightarrow{} \text{independent}$$

$$\underset{\underset{E[g(X)]}{\parallel}}{\int g(x)\,f_x(x)\,dx}\quad \underset{\underset{E[h(Y)]}{\swarrow}}{\int h(y)\,f_y(y)\,dy}$$

Covariance

$$\mathrm{Cov}(X,Y) := E[(X - E(X))(Y - E(Y))]$$

$$E[XY - XE[Y] - YE[X] + E[X]E[Y]]$$

$$= E[XY] - E[X]E[Y] - E[X]\cancel{E[Y]} + \cancel{E[X]}\,E[Y]$$

$$= E[XY] - E[X]E[Y]$$

---

if X & Y are independent
$\mathrm{Cov}(X,Y) = 0$

---

if $\mathrm{Cov}(X,Y) = 0$, are X&Y independent? NO e.g.

$$X = \left\{ \begin{array}{cc} & p \\ 0 & 1/3 \\ 1 & 1/3 \\ -1 & 1/3 \end{array} \right| \quad E[X] = 0$$

$$Y = \begin{cases} 0 & \text{if } x \neq 0 \\ 1 & \text{if } x = 0 \end{cases}$$

$$E[Y] = 1 \cdot \frac{1}{3} = \frac{1}{3}$$

$X \cdot Y = 0$ always

$$E[XY] = 0$$

$$\text{Cov}(X,Y) = E[XY] - E[X]E[Y]$$

$$0 = 0 - 0 \cdot \frac{1}{3}$$

Properties of Covariance.

i) $X = Y$

$$\mathrm{Cov}(X, X) = E[X.X] - E[X]\,E[X]$$

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

$$= \mathrm{Var}(X)$$

ii) $\mathrm{Cov}(X, Y) = \mathrm{Cov}(YX)$

iii) $\mathrm{Cov}(aX, Y) = a.\mathrm{Cov}(X, Y)$

$$\mathrm{Cov}(aX, aX) = a^2\,\mathrm{Cov}(X, X)$$

$$\mathrm{Var}(aX) = a^2\,\mathrm{Var}(X)$$

iv) Covariance is additive

$$\mathrm{Cov}\left(\sum X_i, \sum Y_j\right) = \sum_i \sum_j \mathrm{Cov}(X_i, Y_i)$$

Variance of a sum of random variables.

$$Var\left(\sum_{i=1}^{n} X_i\right) =$$

$$Cov\left(\sum_{i=1}^{n} x_i, \sum_{j=1}^{n} x_j\right)$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} Cov(X_i, X_j)$$

$$= \sum_{i=j}\sum Cov(X_i, X_j) +$$

$$\sum_{i\neq j}\sum Cov(X_i, X_j)$$

$$= \boxed{\sum_{i=1}^{n} Var(X_i) + 2\sum_{i<j}\sum Cov(X_i, X_j)}$$

$$\|$$

$$Var\left(\sum_{i=1}^{n} X_i\right)$$

Binomial R.V.

(n,p) $\quad \mathrm{Var}(X) = np(1-p)$

↳ To prove

$$\begin{matrix} X_1 \\ \vdots \\ X_n \end{matrix} \begin{cases} 1 & \text{if } i^{th} \text{ trial is a success} \\ 0 & \text{if } i^{th} \text{ trial is a failure} \end{cases}$$

$$X = \text{no. of successes} = \sum_{i=1}^{n} X_i$$

$$\mathrm{Var}(X) = \mathrm{Var}\left(\sum_{i=1}^{n} X_i\right)$$

$$= \sum_{i=1}^{n} \mathrm{Var}(X_i) + 2 \sum\sum_{i<j} \underset{\substack{\downarrow \\ 0?}}{\mathrm{Cov}(X_i, X_j)}$$

p

$$X_i = \begin{cases} 1 & p \\ 0 & 1-p \end{cases}$$

$$\mathrm{Var}(X_i) = E[X_i^2] - (E[X_i])^2$$

$$\begin{aligned} E[X_i^2] &= 1^2\ P(X_i = 1) + = p \\ &\quad 0^2\ P(X_i = 0) \end{aligned}$$

$$E[X_i] = 1\ P(X_i = 1) + 0\ P(X_i = 0) = p$$

$$\mathrm{Var}(X_i) = p - p^2 = \underline{p(1-p)}$$

$$\mathrm{Var}(X) = \sum_{i=1}^{n} \mathrm{Var}(X_i) = np(1-p)$$