Fri 2 Feb   24 18 78

Recap

Probability space $(A, \mathcal{F}, P)$

Plan

Law of large numbers

Central limit theorem (probability distributions)

Random walks

Generic probability example

Finite $\Omega = \{w_1, w_2 \cdots w_N\}$ N elements

Measurement X can give same outcome for different states

$$X(w_i) = X(w_j) \quad w_i \neq w_j$$

$$A = \{\underbrace{X_1, X_2, \cdots X_n}_{\text{all distinct}}\} \quad n \text{ elements} \quad n \leq N$$

$$P(X_i \text{ or } X_j) = P(X_i) + P(X_j) = P_i + P_j \quad i \neq j$$

Four fair coin flips

$A = \{HHHH, HHHT, \dots TTTT\}$ $p = 1/16$

$\mathcal{F} = \{$equal H/T, different H/T$\}$

equal H/T: HTHT, HHTT, HTTH, THTH, TTHH, THHT

$P_{equa} = 6/16 = 3/8$ $P_{diff} = 1 - \frac{3}{8} = \frac{5}{8}$

Modelling: Assign probabilities to events

Can be set by symmetries — $\boxed{p = 1/6 \text{ for fair die}}$ page 12

More general: Data-driven modelling

Repeat experiment many time

Monitor outcomes X_i

Infer probabilities p_i

justified by law of large numbers

Expectation value

Consider generic $\mathcal{F} = A = \{X_1, X_2, \cdots X_n\}$ $\sum_{i=1}^{n} p_i = 1$

Useful notation: $\sum_{X \in A} P(X) = 1$

Define expectation value (linear operation)

$$\sum_{X \in A} F(X)\, P(X) = \langle F(X) \rangle$$

Mean of prob. space $\mu = \langle X \rangle = \sum_{X \in A} X\, P(X)$

Variance $\sigma^2 = \langle (X-\mu)^2 \rangle = \sum_{X \in A} (X-\mu)^2 P(X)$

scale of fluctuations around mean

Standard deviation $\sigma = \sqrt{\sigma^2}$ (of prob. space)

$$\sigma^2 = \langle (X - \langle X \rangle)^2 \rangle = \langle X^2 - 2X\langle X\rangle + \langle X\rangle^2 \rangle$$
$$= \langle X^2 \rangle - \cancel{2}\langle X\rangle\langle X\rangle + \langle \cancel{X} \rangle^2$$
$$= \langle X^2 \rangle - \langle X \rangle^2 \qquad \sigma = \sqrt{\langle X^2 \rangle - \langle X \rangle^2}$$

page 14

Repitition

New experiment: Repeat $\mathcal{E}$ & X R times

$\hookrightarrow A = \{X_1, X_2 \dots X_n\}$ n elements

$\downarrow$

outcome space B

For $R=4$ $B = \{X_1 X_1 X_1 X_1,\ X_1 X_2 X_{29} X_1,\ \dots\}$

\# of elements in B? $n \cdot n \cdot n \cdot n \dots = n^R$

$X_i X_j X_k \dots X_z \in B$ built from R $X^{(r)} \in A$ $r = 1, \dots R$

$$P_B(X_i X_j X_k \dots X_z) = P_A(X_i) \cdot P_A(X_j) \cdot P_A(X_k) \dots P_A(X_z)$$

Random variable of repeated experiment:

average $\frac{1}{R} \sum_{r=1}^{R} X^{(r)} = \bar{X}_R$ (arithmetic mean)

Relate to single-experiment $\mu = \langle X^{(r)} \rangle$

assuming μ, σ^2 finite

$$\langle (\bar{X}_R - \mu)^2 \rangle = \left\langle \left(\frac{1}{R} \sum_r X^{(r)} - \mu \right)^2 \right\rangle \qquad \frac{1}{R} \sum_r \mu = \mu$$

$$= \frac{1}{R^2} \left\langle \left(\sum_r (X^{(r)} - \mu) \right)^2 \right\rangle$$

$$= \frac{1}{R^2} \left\langle \left(\sum_r (X^{(r)} - \mu) \right) \cdot \left(\sum_s (X^{(s)} - \mu) \right) \right\rangle$$

$$\left(\sum_i a_i \right) \left(\sum_j b_j \right) = \sum_{i,j} a_i b_j$$

$$= \frac{1}{R^2} \sum_{i,j} \underbrace{\langle (X^{(i)} - \mu)(X^{(j)} - \mu) \rangle}_{\sigma^2 \delta_{ij}}$$

$$\boxed{\langle (\bar{X}_R - \mu)^2 \rangle = \frac{\sigma^2}{R^2} \sum_{i,j} \delta_{ij} = \frac{\sigma^2}{R^2} \underbrace{\sum_i 1}_{R} = \frac{\sigma^2}{R}}$$

$$\lim_{R\to\infty} \langle (\bar{X}_R - \mu)^2 \rangle = 0$$

vanishing sum of squares

$\longrightarrow$ every term zero

$$\lim_{R\to\infty} \bar{X}_R = \mu$$

law of large numbers

$\bar{X}_R \approx \mu$ for $R \gg 1$

Generalize to continuous X with uncountable A

To get probability must integrate over prob. distribution (or density)

$$P(a \leq X \leq b) = \int_a^b p(x)\, dx \qquad x \in \mathbb{R}$$

Generalize expectation values

discrete sums $\longrightarrow$ continuous integrals

$$\langle f(x) \rangle = \int f(x)\, p(x)\, dx$$

(entire domain)

$$\int p(x)\, dx = 1$$

$$\langle x^l \rangle = \int x^l\, p(x)\, dx$$

$l=1 \rightarrow$ mean $\mu = \langle x \rangle$

variance $\sigma^2 = \langle x^2 \rangle - \langle x \rangle^2$ (μ)

$l=2$

Central limit theorem (CLT)

Consider N random variables $x_1, x_2, \cdots x_N$

all with same (finite) μ, σ^2

("i.i.d." = "identical and identically distributed")

Another collective random variable

sum $s = \sum_{i=1}^{N} x_i$

CLT: For $N \gg 1$, prob. distribution $p(s)$ becomes gaussian

$$p(s) \approx \frac{1}{\sqrt{2\pi N\sigma^2}} \exp\left[-\frac{(s-N\mu)^2}{2N\sigma^2}\right]$$

$\int p(s)\, ds = 1$

Collective behaviour of many particles governed by single-particle μ, σ^2

How large is large enough? (N)

CLT application: Random walk

General modelling tool

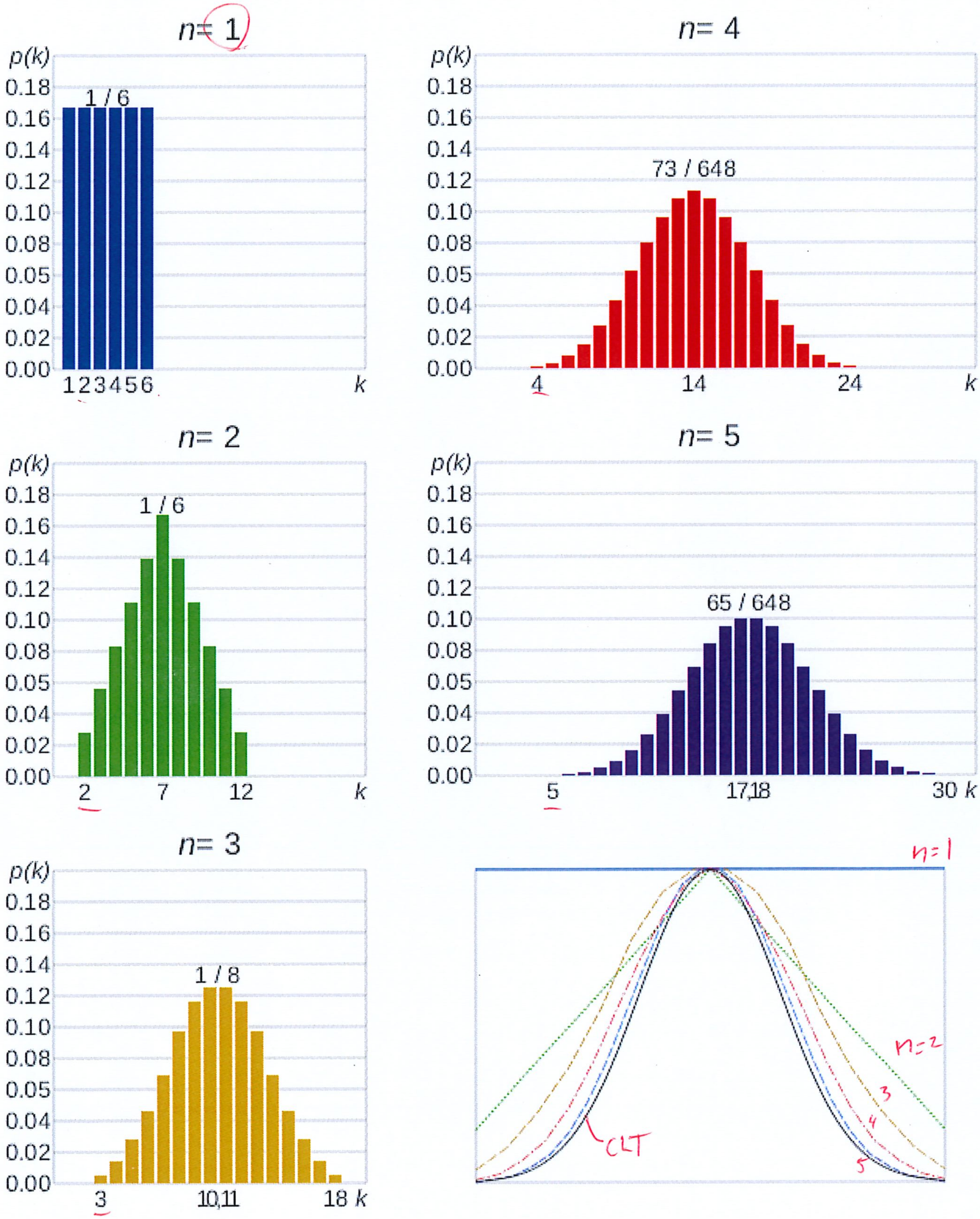
Roulette as random walk in "money space"

Brownian motion, genetic drift, stock prices, ...

Idea: Object takes random step, repeat many times

current state $\rightarrow$ new state

"Markov process" produces Markov chain



commons.wikimedia.org/wiki/File:Dice_sum_central_limit_theorem.svg

Simple example

Step only right $(+\ell)$ or left $(-\ell)$ along line
prob. p — prob. $q = 1-p$

Set $\ell = 1$ and step every Δt (regular time interval)
for N steps $\rightarrow t = N\Delta t$ total time

Representative walk: LRLRRR — $N=6$
start $x_0 = 0$, final $x = 2$

RLRLLL $\rightarrow x = -2$

How many walks with $N=6$?
$$\boxed{2 \cdot 2 \cdot 2 \cdots 2 = 2^N = 2^6 = 64}$$ page 18

$P(\text{LRLRRR}) = \frac{1}{64}$ iff $p = q = \frac{1}{2}$
$$\boxed{qpqppp = p^4 q^2}$$ more generally

page 18

Given an N-step walk with r steps to the right
$N-r$ steps to the left

$$\sim \boxed{p^r q^{N-r}}$$ page 18

Could be many N-step walks with r to right
order doesn't matter

How many ways to get $x = 4$ with $N = 6$?
RRRRRL, LRRRRR, RLRRRR

$$6 = \binom{6}{5} = \binom{N}{r}$$

Overall $\underline{P_r} = \binom{N}{r} p^r q^{N-r}$

Compute expected final position $\langle x \rangle$ and variance $\langle x^2 \rangle - \langle x \rangle^2$
after N steps

$\langle x \rangle = \sum_x x P(x) = \sum_r (2r - N) P_r$

page 19 $$\boxed{x = (+1) r + (-1)(N-r) = 2r - N}$$