23/11 # LECTURE 15 - THE RISE OF URNS.

Eg 15.1 Let $(\Omega_i, \mathbb{P}_i)$, $i = 1, \ldots, n$ be prob. spaces. Let $\tilde{p}$ be a pmf on $[n]$. We select i at random according to $\tilde{p}$ & then choose an element of Ω_i at random according to $\mathbb{P}_i$. What is the prob. space?

Eg 15.2. There are $(m+1)$ urns labelled $0, \ldots, m$. Urn k contains k red balls & $m-k$ white balls. An urn is chosen at random & n balls are sampled from the urn with replacement.

$n \geq 1$ Given that n balls are red, what is the prob. $(n+1)^{th}$ ball is also red?

Soln. $A_n = \{$all 'n' balls are red$\}$ $A_{n+1} = \{ \ldots \}$

Qn. is $P(A_{n+1} \mid A_n) = ?$

$$\mathbb{P}(A_{n+1} \mid A_n) \underset{\text{defn}}{=} \frac{\mathbb{P}(A_{n+1} \cap A_n)}{\mathbb{P}(A_n)} = \frac{\mathbb{P}(A_{n+1})}{\mathbb{P}(A_n)}.$$

$B_k = \{\text{Urn 'k' is chosen}\}$; By assumption $\mathbb{P}(B_k) = \frac{1}{m+1}$.

$\mathbb{P}(A_n | B_k) = \left(\frac{k}{m}\right)^n$ (prob. of choosing red ball in k^{th} urn is k/m)

LTP $\Rightarrow$ $\mathbb{P}(A_n) = \sum_{k=0}^{m} \mathbb{P}(A_n | B_k)\mathbb{P}(B_k)$

$$= \frac{1}{m+1} \sum_{k=1}^{m} \left(\frac{k}{m}\right)^n \xrightarrow[m \to \infty]{\text{Ex.}} \frac{1}{n+1}$$

So $\mathbb{P}(A_{n+1} | A_n) \approx \frac{n+1}{n+2}$. [Laplace's rule of succession]

$(0, \cdots, m)$

Laplace Sun-rise prediction: Each urn is a universe. In k^{th} universe prob. of sun rising is k/m. Prob. (sun will rise tmrw | it has risen 80 year) $= \cdots$

$n = 1,826,213$.

Eg 15.3 URN MODELS :

There is an urn that contains $r \geq 0$ red balls & $b \geq 0$ black balls. A ball is selected at random & put back after noting the colour.

- 'c' new balls of the same colour are added. $c \in \mathbb{Z}$. } fixed.
- 'd' new balls of the opposite colour are added. $d \in \mathbb{Z}$. }

The urn contains $r+b+c+d$ balls, & we repeat this procedure for n steps as long as # of balls ≥ 0. (if $r+c \geq 0$, $b+d \geq 0$.)

$c < 0 \Rightarrow -c$ balls are removed. Same for d.

Very hard to analyze except special cases. For more, Antar B (ISI-D) & Krishanu M. (ISI-K).

Case 1 : $c = -1$, $d = 0$. — At each step a ball is removed at random.

Process will go on until $r+b$ steps.

B_i — black ball is chosen at i^{th} step ; R_i - Red ball is chosen at i^{th} step

$$\mathbb{P}(B_1) = \frac{b}{r+b} = 1 - \mathbb{P}(R_1).$$

$$\left.\begin{aligned} \mathbb{P}(B_2 \mid B_1) &= \frac{b-1}{r+b-1} \\ \mathbb{P}(B_2 \mid R_1) &= \frac{b}{r+b-1} \end{aligned}\right\} \text{ use LTP to compute } \mathbb{P}(B_2).$$

$$\mathbb{P}(B_3 \mid B_2 \cap B_1) = \frac{b-2}{r+b-2} \qquad \mathbb{P}(B_3 \mid R_1 \cap B_2) = \frac{b-1}{r+b-2}$$

$$\mathbb{P}(B_3 \mid R_1 \cap R_2) = \frac{b}{r+b-2}. \qquad \mathbb{P}(B_3 \mid B_1 \cap R_2) = \frac{b-1}{r+b-2}$$

$$\mathbb{P}(B_3 \cap R_2 \cap B_1) = \frac{b-1}{r+b-2}\,\mathbb{P}(R_2 \cap B_1)$$

$$= \frac{b-1}{r+b-2}\,\frac{r}{r+b-1}\cdot\frac{b}{r+b}. = \frac{b(b-1)\,r}{(r+b)(r+b-1)(r+b-2)}$$

$$\mathbb{P}(B_3 \cap B_2 \cap R_1) = \frac{b-1}{r+b-2}\cdot\frac{b}{r+b-1}\cdot\frac{r}{r+b} = \downarrow.$$

$\Rightarrow$ order doesn't seem to matter. Can you prove it? EX.

Ex. 15.4. Suppose $r \geq b$. What is the prob. that # of red balls removed upto step k $\geq$ # of black balls removed upto step k. $\forall k = 1, \ldots, r+b$.

[Ballot problem]

Eg 15.5 (POLYA'S URN) $d = 0$; $c > 0$.

At each stage, more balls of chosen colour are added. [Reinforcement model]

Analyse as before $\mathbb{P}(B_2 \cap B_1) = \dfrac{b(b+c)}{(b+r)(b+r+c)}$ $\quad \mathbb{P}(B_2 \cap R_1) = \dfrac{br}{(b+r)(b+r+c)}$

$$\mathbb{P}(B_2 \cap R_1) = \mathbb{P}(R_2 \cap B_1)$$

$$\mathbb{P}(R_2 \cap R_1) = \frac{r(r+c)}{(b+r)(b+r+c)}$$

$n \geq 3$ $\quad S \subseteq [n]$ $\quad |S| = n_1, \; n - n_1 = n_2$: $A_S = \bigcap_{i \in S} R_i \cap \bigcap_{j \in S^c} B_j$.

red balls chosen at indices in S & Black balls chosen at indices in S^c.

$n \geq 3$

$$\mathbb{P}(A_S) = \begin{cases} \dfrac{b(b+c)\cdots(b+nc-c)}{(b+r)(b+r+c)\cdots(b+r+(n-1)c)} & n_1 = 0 \\[2ex] \dfrac{r(r+c)\cdots(r+(n-1)c)}{(b+r)\cdots\cdots(b+r+(n-1)c)} & n_2 = 0 \\[2ex] \dfrac{b(b+c)\cdots(b+(n_2-1)c)\, r(r+c)\cdots(r+(n_1-1)c)}{(b+r)\cdots\cdots(b+r+(n-1)c)} & n_1 > 0,\ n_2 > 0 \end{cases}$$

Proof: By induction on n.
True for $n = 1, 2$. Assume $n_1 > 0$, $n_2 > 0$ [Formula is true for $n_1 = 0$ or $n_2 = 0$]

WLOG, let $n \in S^c$. Set $B = \bigcap_{i \in S} R_i \cap \bigcap_{j \in S^c \setminus n} B_j$

Since $A_S = B \cap B_n$

$$\mathbb{P}(A_S) = \mathbb{P}(B_n | B)\, \mathbb{P}(B).$$

$$= \frac{b+(n_2-1)c}{b+r+(n-1)c} \mathbb{P}(B)$$

If B happens, we have chosen n_1 red balls & n_2-1 black balls in the first $(n-1)$ steps $\Rightarrow$ $n_1 c$ new red balls are added & $(n_2-1)c$ new black balls are added.

$$\mathbb{P}(B_n \mid B) = \frac{\#\text{ black balls given } B}{\#\text{ balls given } B} = \frac{b+(n_2-1)c}{b+r+(n-1)c}.$$

By induction hypothesis,

$$\mathbb{P}(B) = \frac{b(b+c)\cdots(b+(n_2-2)c)\, r(r+c)\cdots(r+(n_1-1)c)}{(b+r)\cdots(b+r+(n-2)c)}$$

substitute formulas for $\mathbb{P}(B_n|B)$ & $\mathbb{P}(B)$ to derive formula for $\mathbb{P}(A_S)$. $\therefore$ Wly prove for $n \in S$.

Eg 15.6 (Ehrenfest urn model). Two chambers & n particles in total.

A particle is chosen at random & moved to the other chamber.

Colour particles in chamber I in Red ; chamber II in black.

whenever a particle is chosen, change its colour.

This is urn model with $c=-1$, $d=1$. # of balls $= r+b = N$

$$NR_{k,l} = \{\# \text{ of red balls after } l \text{ rounds} = k\} \quad 0 \le k \le N$$

$$= NB_{N-k,l}$$

$$\mathbb{P}(R_{l+1} \mid NR_{k,l}) = \frac{k}{N} = 1 - \mathbb{P}(B_{l+1} \mid NR_{k,l})$$

(LTP) $$\mathbb{P}(NR_{k,l}) = \sum_{j=0}^{N} \mathbb{P}(NR_{k,l} \mid NR_{j,l-1}) \, \mathbb{P}(NR_{j,l-1})$$

At each step # of red balls $\nearrow +1$, $\searrow -1$; $\mathbb{P}(NR_{k,l} \mid NR_{j,l-1}) = 0$ if $j \neq k+1, k-1$

$$= \mathbb{P}(NR_{k,l} \mid NR_{k+1,l-1})\, \mathbb{P}(NR_{k+1,l-1}) + \mathbb{P}(NR_{k,l} \mid NR_{k-1,l-1})\, \mathbb{P}(NR_{k-1,l-1})$$

$$\mathbb{P}(NR_{k,l}) = \frac{k+1}{N}\, \mathbb{P}(NR_{k+1,l-1}) + \frac{N-k+1}{N}\, \mathbb{P}(NR_{k-1,l-1})$$

[Since $\mathbb{P}(NR_{k,l} \mid NR_{k+1,l-1}) = \mathbb{P}(R_l \mid NR_{k+1,l-1}) = \frac{k+1}{N}$

$\mathbb{P}(NR_{k,l} \mid NR_{k-1,l-1}) = \mathbb{P}(B_l \mid NR_{k-1,l-1}) = \frac{N-k+1}{N}$]

Assume $\mathbb{P}(NR_{k,l}) \longrightarrow u_k$ as $l \to \infty$ $\forall\ 0 \le k \le N$.

Then we get that $u_k = \frac{k+1}{N} u_{k+1} + \frac{N-k+1}{N} u_{k-1}$; $u_{-1} = u_{N+1} = 0$.

Solve this i.e., compute u_k.