7/10 LECTURE 5: MORE EXAMPLES ON "EQUALLY LIKELY OUTCOMES".

For us prob. space, finite set Ω & pmf p or PD $\mathbb{P}$.

"In school" — finite set Ω & $P(A) = \frac{|A|}{|\Omega|} = \frac{\#\text{ favourable outcomes}}{\#\text{ outcomes}}$.

$P(A)$ || Prob. of A

Such a P as above is what we call as uniform PD or "Equally likely outcomes".

For eg. by saying Ω is a prob. space with equally likely outcomes means that the PD $\mathbb{P}$ on Ω is uniform PD.

Why go from "Eq. lik. outcomes" to more gen. prob. spaces?

- In reality, not all outcomes are equally likely.

Eg 5.1: $\Omega = \{W, L\}$ if "eq. lik. out." then $p(W) = p(L) = 1/2$.

Eg 5.2 [Sampling with replacement].

A box contains 'n' labelled objects. We pick an object, note down the label & put it back. Then we repeat this process (k-1) times more.

- Choosing k objects out of n (with repetition/replacement).

Sample space Ω = set of k-tuples from $[n] := \{1, \dots, n\}$

$= \{(r_1, \dots, r_k) : r_i \in [n] \ \forall i\}$

['k' labelled particles into 'n' cells]

$= \underbrace{[n] \times [n] \times \dots [n]}_{k-\text{times}}$ ($A \times B$ – Cartesian product)

[$A \times B = \{(a,b) : a \in A, b \in B\}$, A & B are sets]

$=: [n]^k$ ($A^k = \underbrace{A \times \dots \times A}_{k-\text{times}}$)

$A^k = \{(a_1, \dots, a_k) : a_i \in A \ \forall i\}$

Equally likely outcomes $\to$ $p(\omega) = \frac{1}{|\Omega|} = \frac{1}{n^k}$, $\omega \in \Omega$

$\omega = (\omega_1, \dots, \omega_k)$.

(Ω, p) or $(\Omega, \mathbb{P})$ is said as "sampling with replacement from $[n]$ k- many times".

Eg 5.3: [Hashing] There are k individuals & each individual is assigned a code. There are n distinct codes/labels.

For proper identification, codes of k ind. "should" be distinct.

Keeping track of prev. codes isn't easy.

Say a 'lazy' person assigns (hash) codes randomly.

Sample space = set of all possible assignment of (hash) codes for k individuals.

$$= \{(h_1, \dots, h_k) : h_i \in [n] \; \forall i\}$$

$$= [n]^k$$

h_i — hash code of i^{th} person.

All assignments are equally likely.

$$\Rightarrow \quad p(\omega) = \frac{1}{n^k} \qquad \omega = (\omega_1, \dots, \omega_k) \in [n]^k$$

(Error) $E = \{h_1, \dots, h_k\text{'s are not distinct}\}$

(Good) $G = (E)^c = \{h_i\text{'s are distinct}\}$

$$= \{(h_1, \dots, h_k) \in \Omega : h_i \neq h_j \; \forall i \neq j\}$$

Prob. (lazy hash code is good) $= \mathbb{P}(G)$ (defn of G)

$$= \frac{|G|}{|\Omega|} \quad \text{(Eq. likely outcomes)}$$

$$= \frac{|G|}{n^k}$$

$$|G| = n(n-1) \cdots (n-k+1) = \prod_{i=0}^{k-1} (n-i)$$

$$\mathbb{P}(G) = \prod_{i=0}^{k-1} \left(1 - \frac{i}{n}\right)$$

Eg 5.4 (Birthday problem) — If there are 'k' people in a room what is the chance a pair shares same birthday? [Von Mises & others]

Label days as $1, \ldots, 365$; Set $n = 365$.

b_i = birthday of i^{th} person.

Ω = all possible birthdays of 'k' people

$= \{(b_1, \ldots, b_k) : b_i \in [n] \ \forall i\}$ (Same as sampling with repl., or hashing)

Assume any ordering of b_i's is equally likely.

$$P(w) = \frac{1}{n^k} \qquad w = (w_1, \ldots, w_k) \in \Omega.$$

$\mathbb{P}(\exists$ at least one pair sharing same birthday$) = \mathbb{P}(\{(b_1, \ldots, b_k) : b_i = b_j$ for some $i \neq j\})$

(from computation of $\mathbb{P}(B_i)$ in Eg 5.3) $= 1 - \prod_{i=0}^{k-1}\left(1 - \frac{i}{n}\right) = c(k)$.
$n = 365$

$c(2) = 0.002$, $c(5) = 0.027$, $c(10) = 0.116$, $c(23) = 0.507$

If there are 23 people, $\mathbb{P}$(at least 2 people share same birthday$) > 1/2$.

$c(40) > 0.8$ $\qquad c(50) \approx 1$.

$\mathbb{P}$(person 1 shares a birthday with one another)

$= 1 - \mathbb{P}$(person 1 doesn't share a birthday with anybody)

$= 1 - \mathbb{P}(\{(b_1, \ldots, b_k) : b_j \neq b_1 \ \forall j \neq 1\})$

$$= 1 - \frac{|\{(b_1, \ldots, b_k) : b_j \neq b_1 \ \forall j \neq 1\}|}{n^k}$$

$$= 1 - \frac{\underbrace{n}_{b_1} \cdot (n-1)^{k-1}}{n^k} = 1 - \left(1 - \frac{1}{n}\right)^{k-1} = b(k)$$

$b(40) \lesssim 0.11$, $c(40) \gtrsim 0.89$.

$\approx \frac{40}{365}$

$\boxed{\text{THM } 4.1}$ Let $(\Omega, \mathbb{P})$ be a fin. prob. space. Then the following hold.

(i) $\mathbb{P}(A) \leq \mathbb{P}(B) \quad \forall \quad A \subseteq B \subseteq \Omega$. [Monotonicity]

(ii) $\mathbb{P}(A) \leq 1 \quad \forall \quad A \subseteq \Omega$

(iii) $\mathbb{P}(\phi) = 0$

(iv) $\mathbb{P}(A_1 \cup \cdots \cup A_n) \leq \sum_{i=1}^{n} \mathbb{P}(A_i) \quad \forall \; A_1, \ldots, A_n \subseteq \Omega$. [finite subadditivity]

(v) $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$ [Inclusion-Exclusion]

Proof: (i) $\mathbb{P}(A) = \sum_{\omega \in A} p(\omega) \leq \sum_{\omega \in B} p(\omega)$ $(A \subseteq B \text{ \& } p(\omega) \geq 0 \; \forall \omega)$

$= \mathbb{P}(B)$ (defn)

(ii) Set $B = \Omega$. (i) $\Rightarrow$ $\mathbb{P}(A) \leq \mathbb{P}(\Omega) = 1$

(prob. of $\mathbb{P}$)

(iii) $\mathbb{P}(A \cup \phi) = \mathbb{P}(A) + \mathbb{P}(\phi)$ as $A \cap \phi = \phi$.

$\Rightarrow \mathbb{P}(A) = \mathbb{P}(A) + \mathbb{P}(\phi)$ as $A \cup \phi = A$.

so $\mathbb{P}(\phi) = 0$.

(iv) we'll prove by induction.

$n = 1$ is trivially true. $\mathbb{P}(A_1) \leq \mathbb{P}(A_1)$ ✓

Let $n = 2$. Then $A_1 \cup A_2 = A_1 \sqcup (A_2 \setminus A_1)$

disjoint union

$\mathbb{P}(A_1 \cup A_2) = \mathbb{P}(A_1 \sqcup (A_2 \setminus A_1))$

$= \mathbb{P}(A_1) + \mathbb{P}(A_2 \setminus A_1)$ (finite additivity of $\mathbb{P}$) $(A_1 \cap (A_2 \setminus A_1) = \phi)$

$\leq \mathbb{P}(A_1) + \mathbb{P}(A_2)$

(monotonicity as $A_2 \setminus A_1 \subseteq A_2$)

Assume finite subadd. holds for $n-1 \geq 2$.
Consider $A_1, \ldots, A_n$. We've to prove fin. subadd for n.

$$A_1 \cup \ldots \cup A_n = A_1 \sqcup \left((A_2 \cup \ldots \cup A_n) / A_1 \right)$$

$$\begin{aligned}
\mathbb{P}(A_1 \cup \ldots \cup A_n) &= \mathbb{P}(A_1) + \mathbb{P}((A_2 \cup \ldots \cup A_n)/A_1) \quad (\text{fin. add.}) \\
&\leq \mathbb{P}(A_1) + \mathbb{P}(A_2 \cup \ldots \cup A_n) \quad (\text{mon.}) \\
&\leq \mathbb{P}(A_1) + \mathbb{P}(A_2) + \cdots + \mathbb{P}(A_n) \quad (\text{fin. subadd. holds for } n-1)
\end{aligned}$$

(v) $$A \cup B = A \sqcup (B / A \cap B)$$

$$\begin{aligned}
\mathbb{P}(A \cup B) &= \mathbb{P}(A) + \mathbb{P}(B / A \cap B) \quad (\text{fin. add.}) \\
&= \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B) \quad (A \cap B \subseteq B \text{ \& use Lemma 4.4})
\end{aligned}$$

$\blacksquare$