02 February 2023

§1.5. Discrete Probability

Randomness is useful: e.g. games (like dice, lotto, stock markets...)

Probability space: (Ω, P) — P: probability, Ω: finite set

Normalized: $\sum_{\omega \in \Omega} P(\omega) = 1$.

E.g.: Throw a coin $\Rightarrow \Omega = \{\text{Head, Tail}\}$, $|\Omega| = 2$

If the coin is fair, then we have $P(\text{Head}) = P(\text{Tail}) = \frac{1}{2}$.

I.e. if we throw a coin many times, then approximately $\frac{1}{2}$ of the times we see Head, $\frac{1}{2}$ of the times we see Tail.

E.g.: Throw a 6-side dice.

$\Omega = \{1, 2, 3, 4, 5, 6\}$

Fair dice: $P_0 = \{\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}\}$

An unfair dice: $P_1 = \{\frac{1}{4}, \frac{1}{10}, \frac{1}{5}, \frac{1}{4}, \frac{1}{10}, \frac{1}{10}\}$.

Event: $\emptyset \neq A \subseteq \Omega$

$P(A) = \sum_{\omega \in A} P(\omega)$.

Example: For a fair dice (P_0 above)

$A_1 = \{\text{number is odd}\} = \{1, 3, 5\}$

$A_2 = \{\text{number is even}\} = \{2, 4, 6\}$

$A_3 = \{\text{number is} \geq 2\} = \{2, 3, 4, 5, 6\}$

$P(A_1) = \sum_{\omega \in A_1} P(\omega) = \frac{3}{6}$

$P(A_2) = \frac{3}{6}$

$P(A_3) = \frac{5}{6}$.

Proposition: a) $A \subseteq B \Rightarrow P(A) \leq P(B)$.

b) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

c) If $A = \coprod_i A_i \Rightarrow P(A) = \sum P(A_i)$

d) $P(\Omega \setminus A) = 1 - P(A)$

e) $P(\bigcup_j A_i) \leq \sum P(A_i)$.

Proof: Most of them can be proven by Venn's diagram.

a) $P(B) = P(A) + P(B \setminus A)$ (by c) $\geq P(A)$.

b) $A \cup B = (A \setminus B) \amalg (A \cap B) \amalg (B \setminus A)$

$P(A \cup B) = P(A \setminus B) + P(A \cap B) + P(B \setminus A)$ (by c)

$P(A \setminus B) = P(A) - P(A \cap B)$

$A \setminus B$ $A \cap B$ A

$P(B \setminus A) = P(B) - P(A \cap B)$

$B \setminus A$ $A \cap B$ B

c) $P(\coprod A_i) = \sum_{\omega \in \coprod A_i} P(\omega)$

$= \sum_i \left(\sum_{\omega \in A_i} P(\omega) \right) = \sum_i P(A_i)$

d) A $\Omega \setminus A$ Ω Apply c.

e) Proof by induction.

$P(\bigcup_{i=1}^{k} A_i) \leq \sum_{i=1}^{k} P(A_i)$

$k = 2$: $P(A_1 \cup A_2) \overset{\text{by b}}{=} P(A_1) + P(A_2) - P(A_1 \cap A_2) \overset{\geq 0}{\leq} P(A_1) + P(A_2)$ ✓

Assume true for k. Then for $k+1$:

$P(\bigcup_{i=1}^{k+1} A_i) = P((\bigcup_{i=1}^{k} A_i) \cup A_{k+1})$

$\overset{\text{by b}}{=} P(\bigcup_{i=1}^{k} A_i) + P(A_{k+1})$

$\overset{\text{by } k}{\leq} \sum_{i=1}^{k} P(A_i) + P(A_{k+1})$

$= \sum_{i=1}^{k+1} P(A_i)$. □

Now assume that we have two probability spaces (Ω_1, p_1), (Ω_2, p_2).

Then we can consider a new probability $(\Omega_1 \times \Omega_2, p)$.

(ω_1, ω_2)

Need to give a probability $p(\omega_1, \omega_2)$.

It could happen that there is some relation between Ω_1 / Ω_2 which affects p.

Example: like if one person has some serious disease already, then the probability that they will die by covid could be 20%.

For a general person, the probability to die because of Covid can be extremely low, like 0.01%.

The in general p of the product space $\Omega_1 \times \Omega_2$ may be difficult to determine. But there is a nice case:

Independent spaces: (or random variables) Say $(\Omega_1, p_1), (\Omega_2, p_2)$ are independent if $\forall \omega_1 \in \Omega_1, \omega_2 \in \Omega_2$

$p(\omega_1, \omega_2) = p_1(\omega_1) \cdot p_2(\omega_2)$. □

(So we can compute p using p_1 & p_2.)

Example: If we throw a fair dice twice, independently, then

$P(\text{sum of the 2 numbers} = 3) = ?$

$(\Omega_1, P_1) = (\Omega_2, P_2) = (\{1,2,3,4,5,6\}, \{\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}\})$

$\{\text{Sum of the 2 numbers} = 3\} = \{(\omega_1, \omega_2) : \omega_1 \in \Omega_1, \omega_2 \in \Omega_2, \omega_1 + \omega_2 = 3\} \subseteq \Omega_1 \times \Omega_2$

$\Rightarrow \exists \{(1,2), (2,1)\}$.

Since $P(\{\text{sum} = 3\}) = P((1,2)) + P((2,1))$

Since we assume independence:

$P(1,2) = P_1(1) P_2(2) = \frac{1}{6} \times \frac{1}{6} = \frac{1}{36}$

$P(2,1) = P_1(2) P_2(1) = \frac{1}{6} \times \frac{1}{6} = \frac{1}{36}$.

Example: The same, except now the dice is unfair

$(\Omega_1, P_1) = (\Omega_2, P_2) = (\{1,2,\dots,6\}, \{\frac{1}{4}, \frac{1}{10}, \frac{1}{5}, \frac{1}{4}, \frac{1}{10}, \frac{1}{10}\})$

$\sum = 1$

$P(\text{Sum} = 3) = P((1,2)) + P((2,1))$

$= P_1(1) P_2(2) + P_1(2) P_2(1)$

$= \frac{1}{4} \times \frac{1}{10} + \frac{1}{10} \times \frac{1}{4} = \frac{2}{40} = \frac{1}{20}$

Theorem: If the event spaces are independent, then for events $A_1 \subseteq \Omega_1, A_2 \subseteq \Omega_2$:

$P(A_1 \times A_2) = P_1(A_1) \times P_2(A_2)$

(Similarly for more sets).

Proof: $A_1 \times A_2 = \{(\omega_1, \omega_2) \in \Omega_1 \times \Omega_2 : \omega_1 \in A_1, \omega_2 \in A_2\}$

$$P(A_1 \times A_2) = \sum_{\omega_1 \in A_1} \sum_{\omega_2 \in A_2} P(\omega_1, \omega_2)$$

$$\overset{\text{independence}}{=} \sum_{\omega_1 \in A_1} \sum_{\omega_2 \in A_2} P_1(\omega_1) P_2(\omega_2)$$

$$= \sum_{\omega_1 \in A_1} P_1(\omega_1) \underbrace{\left(\sum_{\omega_2 \in A_2} P_2(\omega_2) \right)}_{P_2(A_2)}$$

$$= P_1(A_1) P_2(A_2) \quad \square$$

Random variable: (Ω, P) a probability space & $X : \Omega \to \mathbb{R}$ is called a random variable.

Example: (Ω = lotto numbers, P = uniform)

$X : \Omega \to \mathbb{R}$

(" 10^6 numbers")

1 lotto 1st price win $\mapsto$ 1 000 000 $

1 lotto 2nd price win $\mapsto$ 500 000 $

the other numbers $\mapsto$ 0

probability that a person win 1000 000 $ =

$P(\text{lotto number} = \text{1st price}) = \frac{1}{10^6}$

prob ... win 500 000 =

$P(\text{lotto number} = \text{2nd price}) = \frac{1}{10^6}$.

$P(\text{a number win at least 1 \$}) = P(\text{win 1st price or 2nd price}) = \frac{2}{10^6}$.

$p(\text{a person doesn't win}) = 1 - \frac{2}{10^6}$. □

If X is a random variable:

$$p(X \in [a,b]) \overset{\text{def}}{=} p(\omega \in \Omega : X(\omega) \in [a,b])$$

in general if $Z \subseteq \mathbb{R}$:

$$p(X \in Z) \overset{\text{def}}{=} p(\omega \in \Omega : X(\omega) \in Z).$$ □

Average / Mean: if we have a set of numbers $a_1, \dots, a_n$. Then the average / mean is:

$$\frac{a_1 + \dots + a_n}{n}.$$

Think about:

$$\frac{a_1 + \dots + a_n}{n} = a_1 \cdot \left(\frac{1}{n}\right) + \dots + a_n \left(\frac{1}{n}\right)$$

a_1: probability for the number to be a_1; $\frac{1}{n}$: probability that the number is a_1.

(there, we have a uniform probability / uniform distribution.)

Another way to think about average/mean.

$$(\Omega, p) = \{\{1, \dots, n\}, \{\tfrac{1}{n}, \dots, \tfrac{1}{n}\}\}$$

uniform distribution

$$X : \Omega \to \mathbb{R}, \quad i \mapsto a_i = X(i)$$

$$\frac{a_1 + \dots + a_n}{n} = \sum_{i \in \Omega} X(i)\, p(i) = E(X)$$

$E(X)$: expectation.

Expectation is good indicator of how much you gain / lose in long term.

Example: For the lotto problem above, say the price of a ticket is 10 $.

$$X : \Omega = \{\text{lotto numbers}\} \to \mathbb{R}$$

$$\text{lotto} \mapsto \text{Win of lotto} - \text{price of the lotto}$$

1st price ticket $\mapsto$ 1 000 000 − 10

2nd price ticket $\mapsto$ 50 000 − 10

no win ticket $\mapsto$ −10

$$E(X) = X(\text{1st prize})\, p(\text{1st prize}) + X(\text{2nd prize})\, p(\text{2nd prize}) + X(\text{no win})\, p(\text{no win})$$

$$= (1\,000\,000 - 10)\frac{1}{10^6} + (50\,000 - 10)\frac{1}{10^6} + (-10)\left(1 - \frac{2}{10^6}\right)$$

< 0 You lose if you keep buying lottos !

In general can define a random variable like: $X: (\Omega, p) \to T$ → Some space (usually $T = \mathbb{R}, \mathbb{Z}, \mathbb{N}, \mathbb{C}, \ldots$)

Independent random variables: (nice case)

$X: \Omega \to T$, $Y: \Omega \to U$ are independent if $\forall\ \omega_1, \omega_2 \in \Omega$:

$p(X = \omega_1 \wedge Y = \omega_2) = p(X=\omega_1)\, p(Y=\omega_2)$

(If $\Omega_1 \to T$, $\Omega_2 \to U$, then use p_1, p_2).

Theorem: if X & Y are independent then $E(XY) = E(X)\,E(Y)$.

Proof:

$$E(XY) = \sum_{\omega \in \Omega} p(\omega)\, X(\omega)\, Y(\omega)$$

$$= \sum_{(x,y) \in T \times U} p(\omega : X(\omega) = x \wedge Y(\omega) = y) \cdot x \cdot y$$

$$= \sum_{(x,y) \in T \times U} p(X(\omega) = x) \cdot p(Y(\omega) = y) \cdot x \cdot y$$

$$= \left(\sum_x p(X(\omega) = x) \cdot x \right) \left(\sum_y p(Y(\omega) = y) \cdot y \right)$$

$$= E(X)\,E(Y). \quad \square$$

One interesting example of non-independent random variables (Math Stackexchange)

(Ω, p) ~~$= \{1,2,3\}$~~

$X, Y: \Omega \to V = \{1,2,3\}$ (joint distribution)

X \ Y	1	2	3
1	3/30	0	6/30
2	6/30	3/30	0
3	1/30	7/30	4/30

Mean

$p(X = 1, Y = 1) = 3/30$

$p(X = 3, Y = 2) = 7/30$

$$p(X=1) = p((X=1, Y=1) \text{ or } (X=1, Y=2) \text{ or } (X=1, Y=3))$$

$$= p(X=1, Y=1) + p(X=1, Y=2) + p(X=1, Y=3)$$

$$= \frac{3}{30} + 0 + \frac{6}{30} = \frac{9}{30} = \text{sum of numbers on 1st row}$$

$$p(Y=2) = \text{sum of numbers on the 2nd column} = 0 + \frac{3}{30} + \frac{7}{30} = \frac{10}{30}$$

$p(X=1, Y=2) = 0 \neq p(X=1)\,p(Y=2)$

X & Y are not independent.

Theorem: $E(XY) = E(X)\,E(Y)$!

$$E(XY) = \sum_{\substack{x\in\{1,\dots,3\}\\ y\in\{1,\dots,3\}}} P(X=x, Y=y)\,xy$$

$$= \frac{3}{30}\times 1\times 1 + 0\times 1\times 2 + \frac{6}{30}\times 1\times 3 + \dots$$

$$E(X) = \sum_{x\in\{1,\dots,3\}} P(X=x)\,x = \frac{9}{30}\times 1 + \frac{9}{30}\times 2 + \frac{12}{30}\times 3$$

$$E(Y) = \sum_{y\in\{1,\dots,3\}} P(Y=y)\,y = \cdots \quad \square$$

Properties: $E(X_1+\cdots+X_k) = E(X_1)+\cdots+E(X_k)$

$E(\alpha X) = \alpha E(X)$ (α: constant)

$E(X+c) = E(X)+c$ (c: constant)

Proof: Easy.

Variation: Variation measures the "average distance to the averaged" expectation

$$VX = E\left((X-EX)^2\right)$$

"average" ... "distance" ... "average"

Proposition: $VX = E(X^2) - (EX)^2$.

Proof: Put $\mu = EX$ = constant!

$(X-EX)^2 = (X-\mu)^2$

expand $X^2 - 2X\mu + \mu^2$. $\square$

Application: π a permutation of $\{1,\dots,n\}$.

A fixed point of π is an element $x \in \{1,\dots,n\}$ such that $\pi(x) = x$.

(Recall: π = a bijective map $\{1,\dots,n\} \circlearrowright$)

It is complicated to know how many fixed points π will have.

Example: $\pi = id = (1 \dots n)$

$\Rightarrow$ π has n fixed points.

$\pi = [2 \dots n]$ cycle

then π has no fixed point

On average, how many fixed points which a permutation has?

Precise statement:

Let $(\Omega, P) = \{$ permutations of $\{1,\dots,n\}\}$, uniform distribution)

$$P(\pi) = \frac{1}{|\Omega|} = \frac{1}{n!}$$

~~Then what is like~~ $X: \Omega \to \{0,\dots,n\}$

$X(\pi)$ = # of fixed points of π

what is $E(X)$?

The answer = 1. i.e. If we choose a random permutation π then "it is likely to have 1 fixed point."

Now define:

$f_i : \Omega \to \{0,1\}$

$f_i(\pi) = \begin{cases} 0 & \text{if } \pi(i) \neq i \\ 1 & \text{if } \pi(i) = i \end{cases}$

$f_i(\pi) = 1 \Leftrightarrow i$ is a fixed point of π.

$X(\pi) =$ # of fixed points of $\pi = \sum_{i=1}^{n} f_i(\pi)$

$\Rightarrow E(X) = E\left(\sum_{i=1}^{n} f_i\right)$

$= \sum_{i=1}^{n} E(f_i)$

$E(f_i) = \sum_{\pi \in \Omega} f_i(\pi) \cdot p(\pi) = \frac{1}{n!} \sum_{\pi \in \Omega} f_i(\pi)$

$= \frac{1}{n!}$

$f_i(\pi) = 1 \Leftrightarrow \pi = \pi' \circ [i]$ cycle decomp.

π' is any permutation of $\{1,\dots,n\} \setminus \{i\}$

$\Rightarrow \sum_{\pi \in \Omega} f_i(\pi) = \sum_{\pi' \text{ permutation of } \{1,\dots,n\} \setminus \{i\}} 1 = (n-1)!$

$\Rightarrow E(X) = \sum_{i=1}^{n} E(f_i) = \sum_{i=1}^{n} \frac{1}{n!} \sum_{\pi \in \Omega} f_i(\pi)$

$= \sum_{i=1}^{n} \frac{1}{n!} \frac{(n-1)!}{(n-1)!} = \sum_{i=1}^{n} \frac{1}{n} = 1$

$VX = E(X^2) - E(X)^2$. ☐

$X^2 = \left(\sum_{i=1}^{n} f_i\right)^2 = \sum_{i=1}^{n} f_i^2 + 2\sum_{i<j} f_i f_j$

$f_i \in \{0,1\} \Rightarrow f_i^2 = f_i$

$E(f_i^2) = E(f_i) = \frac{1}{n}$.

~~$E(f_i f_j) = \sum f_i f_j$~~

$f_i f_j = \begin{cases} 1 & \text{if } \pi(i) = i \ \& \ \pi(j) = j \\ 0 & \text{otherwise} \end{cases}$

~~$E(f_i f_j) = \sum_{\pi \in \Omega}$~~

Now if $\pi(i) = i$, $\pi(j) = j \Rightarrow$

$\pi = \pi'' \circ [i] \circ [j]$

where π'' is any permutation of $\{1,\dots,n\} \setminus \{i,j\}$.

$E(f_i f_j) = \sum_{\pi \in \Omega} f_i f_j(\pi)\, p(\pi)$

$= \frac{1}{n!} \sum_{\pi'} 1 = \frac{(n-2)!}{n!}$

π' permutation of $\{1,\dots,n\} \setminus \{i,j\}$

$VX = \frac{E(X^2)}{2} - \frac{E(X)^2}{1^2} = 1$. ☐