April 28

Wednesday, April 28, 2021 2:56 PM

Comp Sci 212
1. Birthday Paradox
2. Random Variables

Announcements
- Midterm May 3 (make-up May 4)
- No homework this week
- Getting to know you meetings

Birthday Paradox
- In a room of n people, what's the probability that 2 people have the same birthday?

$\Pr[S \setminus E] = 1 - \Pr[E]$ (complement rule)

- View birthdays as a sequence, i^{th} element is birthday of i^{th} person.
- Let D be set of birthdays $|D| = 365$
- Sample space $S = D^n = \underbrace{D \times D \times \cdots \times D}_{n \text{ times}}$

uniform distributions
- $E \subseteq S$, no two birthdays are the same

$\Pr[E] = \dfrac{|E|}{|S|} = \dfrac{|E|}{|D|^n}$ $n=2$, $|E| = |D| \cdot (|D|-1)$

generalized

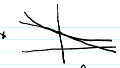
$|E| = |D| \cdot (|D|-1) \cdot (|D|-2) \cdots (|D|-(n-1))$

$\overline{|D| \quad\quad |D| \quad\quad |D| \quad \cdots \quad |D|}$

$= \prod_{i=0}^{n-1} \dfrac{|D|-i}{|D|} = \prod_{i=0}^{n-1} 1 - \dfrac{i}{|D|}$

product from $i=0$ to $i=n-1$

$a \le b$, $-a \ge -b$, $e^{-a} \ge e^{-b}$ $e^a \cdot e^b = e^{a+b}$

$1 - x \approx e^{-x}$  tangent line,

$1 - x \le e^{-x}$ → fact from calculus

$\Pr[E] = \prod_{i=0}^{n-1} 1 - \dfrac{i}{|D|} \le \prod_{i=0}^{n-1} e^{-\frac{i}{|D|}} = e^{-\left(0 + \frac{1}{|D|} + \frac{2}{|D|} + \cdots + \frac{n-1}{|D|}\right)}$

$1 + 2 + \cdots + n-1 = \dfrac{n(n-1)}{2}$

$= e^{-\frac{n(n-1)}{2|D|}}$

$\dfrac{n(n-1)}{2|D|} \ge 1$

If $\dfrac{n(n-1)}{2} \ge |D|$ or, $n \ge \sqrt{2|D|} + 1$,

then probability is $\le e^{-1}$ or at most $\frac{1}{e}$.

$|D| = 365$, $n \approx 27$

Birthday Principle

If there are d days in a year, and $\sqrt{2d}$ people in a room, then the probability that two people share the same birthday is about $1 - \frac{1}{e}$.

S = sample space, uniform distribution

How can we estimate $|S|$ from samples

sequence, each element of sequence random element from S.

- Keep sampling until you think you've seen every element of $S \to |S|$ samples.
- Keep sampling until you see a repeat - after about $\sqrt{2|S|}$ samples.

Def. A random variable is a total function on the sample space of a probability space.

Ex. $S = \{H, T\}^{100}$

$R: S \to \mathbb{R}$, $R(\omega) = \{\#$ of heads in ω

$\Pr[R = 50]$ $\Pr[R \ge 75]$ $\Pr[R \text{ is odd}]$

$E = \{\omega \mid \omega \in S, \omega \text{ has } 50 \text{ heads}\}$

$= \Pr[E]$

uniform dist. $\Pr[R = 50] = \dfrac{|E|}{|S|} = \dfrac{\binom{100}{50}}{2^{100}}$

sticky coin.

$\Pr[\text{all heads}] = \Pr[\text{all tails}] = \frac{1}{2}$ $\Pr[R = 50] = 0.$

$\Pr[R = 100] = \Pr[\text{all heads}] = \frac{1}{2}$

Ex. $M: S \to \mathbb{R}$, $M(\omega) = \begin{cases} 1 & \text{if all coins are same} \\ 0 & \text{o.w.} \end{cases}$

$\Pr[M = 1] = \Pr[E]$ $E = \{\text{all heads, all tails}\}$

uniform - $\dfrac{|E|}{|S|} = \dfrac{2}{2^{100}} = 2^{-99}$

sticky - 1

Def. Two random variables are independent if A, B

$\Pr[A = x_1] \cdot \Pr[B = x_2] = \Pr[A = x_1 \text{ and } B = x_2]$

for all x_1, x_2.

Thm. If uniform distribution, R & M are not independent.

Proof. $\Pr[R = 2 \text{ and } M = 1] = 0$ (no outcomes in both).

$\Pr[M = 1] \ne 0$, $\Pr[R = 2] \ne 0$, product of two is non-zero.

□

Thm. If sticky coin, R & M are independent.

Proof. $\Pr[R = k \text{ and } M = 0] = 0$ for all k

↳ M is never 0

$\Pr[M = 0] = 0$, so $\Pr[R = k] \cdot \Pr[M = 0] = 0$

$\left.\right\} x_2 = 0$

$\Pr[R = k \text{ and } M = 1] = \Pr[R = k] = \Pr[R = k] \cdot \Pr[M = 1]$.

$\sum_{\substack{\omega: R(\omega) = k \\ M(\omega) = 1}} \Pr[\omega] = \sum_{\omega: R(\omega) = k} \Pr[\omega]$

$(\Pr[M = 1] = 1)$ □

$\left.\right\} x_2 = 1$

If $M(\omega) = 0$, then $\Pr[\omega] = 0$

$w: R(w)=k$ $\quad$ $w: R(w)=k$ $\quad$ $(\Pr[M=1]=1)$

$M(w)=1$

If $M(w)=0$, then

$\Pr[w]=0$