

Let  $E_i = A$  wins starting with  $\$i$ .

What is  $P(E_i)$ ? Let  $p_i = P(E_i)$ .

Condition on outcome of the first flip, and use the LTP.

$$\begin{aligned} p_i = P(E_i) &= P(E_i | H)P(H) + P(E_i | T)P(T) \\ &= P(E_{i+1}) \cdot \frac{1}{2} + P(E_{i-1}) \cdot \frac{1}{2} \\ &= \frac{1}{2}(p_{i+1} + p_{i-1}) \end{aligned}$$

$$2p_i = p_{i+1} + p_{i-1}$$

$$p_i - p_{i-1} = p_{i+1} - p_i$$

$$p_2 - p_1 = p_1 - p_0 = p_1 - 0 = p_1$$

$$p_2 = 2p_1$$

$$p_3 = 3p_1$$

$$p_4 = 4p_1$$

$$p_i = ip_1 \text{ for } 0 \leq i \leq N.$$

$$1 = p_N = Np_1$$

$$p_1 = 1/N$$

$$p_i = i/N$$

Bayes' Theorem (Rev. Thomas Bayes, c. 1701-1761):

$$P(F|E) = \frac{P(E|F)P(F)}{P(E)}$$

Proof:  $P(F|E) = \frac{P(E \cap F)}{P(E)} = \frac{P(E|F)P(F)}{P(E)}$

Corollary:  $P(F|E) = \frac{P(E|F)P(F)}{P(E|F)P(F) + P(E|\bar{F})P(\bar{F})}$

Why it's useful:

$$P(\text{disease} | \text{test result}) \sim P(\text{test result} | \text{disease})$$

Ex: 60% of email is spam.

90% of spam has a forged header

20% of nonspam has a forged header.

~~What~~ Let  $F$  = forged header,  $J$  = spam

What is  $P(J|F)$ ?

$$\begin{aligned} P(J|F) &= \frac{P(F|J)P(J)}{P(F|J)P(J) + P(F|\bar{J})P(\bar{J})} \\ &= \frac{0.9 \times 0.6}{0.9 \times 0.6 + 0.2 \times (1-0.6)} \approx 0.871 \end{aligned}$$

"prior"  $P(J) = 0.6$

"posterior"  $P(J|F) \approx 0.871$



Ex. Paternity testing

Child has  $(A, a)$  gene pair: event  $B_{Aa}$   
 Mother has  $(A, A')$

Two possible fathers,  $F_1 = (a, a)$ ,  $F_2 = (A, a)$ .

$$P(F_1) = p, \quad P(F_2) = 1 - p.$$

$$P(F_1 | B_{Aa}) = \frac{P(B_{Aa} | F_1) P(F_1)}{P(B_{Aa} | F_1) P(F_1) + P(B_{Aa} | F_2) P(F_2)}$$

$$= \frac{1 \cdot p}{1 \cdot p + \frac{1}{2}(1-p)} = \frac{2p}{2p + 1 - p} = \frac{2p}{1+p}$$

$$\frac{2p}{1+p} \geq \frac{2p}{1+1} = p. \quad \text{E.g., } p = \frac{1}{2}, \quad \frac{2p}{1+p} = \frac{2}{3}$$