Conditional Probability

P( | )

General defn: $P(E \mid F) = \dfrac{P(EF)}{P(F)}$ where P(F) > 0

Implies: P(EF) = P(E|F) P(F) (*"the chain rule"*)

General definition of Chain Rule:

$$P(E_1 E_2 \cdots E_n) = P(E_1) P(E_2 \mid E_1) P(E_3 \mid E_1, E_2) \cdots P(E_n \mid E_1, E_2, \ldots, E_{n-1})$$

$P(E) = P(EF) + P(EF^c)$
$= P(E|F)\,P(F) + P(E|F^c)\,P(F^c)$
$= P(E|F)\,P(F) + P(E|F^c)\,(1-P(F))$

weighted average, conditioned on event F happening or not.

More generally, if $F_1, F_2, \ldots, F_n$ partition S (mutually exclusive, $\cup_i F_i = S$, $P(F_i)>0$), then

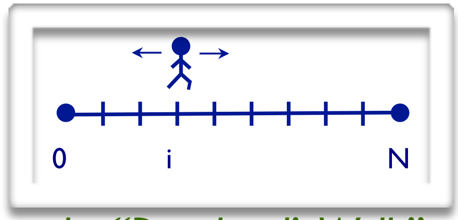
$P(E) = \sum_i P(E|F_i)\,P(F_i)$

weighted average, conditioned on events F_i happening or not.

(Analogous to reasoning by cases; both are very handy.)

Gambler's ruin

BT pg. 63



aka "Drunkard's Walk"

A gambler makes a series of independent bets. In each bet, he wins \$1 with probability ½ and loses \$1 with probability ½. He starts with \$k.
What is the probability he makes \$N before he goes broke?

nice example of the utility of conditioning: future decomposed into two crisp cases instead of being a blurred superposition thereof

$$\begin{aligned}
p_i &= P(E_i) = P(E_i \mid H)P(H) + P(E_i \mid T)P(T) \\
p_i &= \frac{1}{2}(p_{i+1} + p_{i-1}) \\
2p_i &= p_{i+1} + p_{i-1} \\
p_{i+1} - p_i &= p_i - p_{i-1} \\
p_2 - p_1 &= p_1 - p_0 = p_1, \ \text{since } p_0 = 0
\end{aligned}$$

So:

$$\begin{aligned}
p_2 &= 2p_1 \\
&\cdots \\
p_i &= ip_1 \\
p_N &= Np_1 = 1 \\
\boxed{p_i} &\boxed{= i/N}
\end{aligned}$$

Most common form:

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

Expanded form (using law of total probability):

$$P(F \mid E) = \frac{P(E \mid F)P(F)}{P(E \mid F)P(F) + P(E \mid F^c)P(F^c)}$$

Why it's important:

Reverse conditioning

P(model | data) ~ P(data | model)

Combine new evidence (E) with prior belief (P(F))

Posterior vs prior

HIV testing

	HIV+	HIV-
Test +	$0.98 = P(E\|F)$	$0.01 = P(E\|F^c)$
Test -	$0.02 = P(E^c\|F)$	$0.99 = P(E^c\|F^c)$

0.5% of population has HIV

Let E = you test positive for HIV

Let F = you actually have HIV

What is P(F|E) ?

$$\begin{aligned} P(F \mid E) &= \frac{P(E \mid F)P(F)}{P(E \mid F)P(F) + P(E \mid F^c)P(F^c)} \\ &= \frac{(0.98)(0.005)}{(0.98)(0.005) + (0.01)(1 - 0.005)} \\ &\approx 0.330 \end{aligned}$$

↖ $P(E) \approx 1.5\%$

Note difference between conditional and joint probability: P(F|E) = 33% ; P(FE) = 0.49%

	HIV+	HIV-
Test +	0.98 = $P(E \mid F)$	0.01 = $P(E \mid F^c)$
Test -	0.02 = $P(E^c \mid F)$	0.99 = $P(E^c \mid F^c)$

Let E^C = you test **negative** for HIV

Let F = you actually have HIV

What is $P(F \mid E^C)$?

$$\begin{aligned} P(F \mid E^c) &= \frac{P(E^c \mid F)P(F)}{P(E^c \mid F)P(F) + P(E^c \mid F^c)P(F^c)} \\ &= \frac{(0.02)(0.005)}{(0.02)(0.005) + (0.99)(1 - 0.005)} \\ &\approx 0.0001 \end{aligned}$$