Information theory

Methods: Logic, Part 7

Michael Franke

Subjective beliefs & information gained

	sunny	cloudy	rainy
Jones' beliefs	0.6	0.2	0.2
Smith's beliefs	0.1	0.2	0.7

Information content

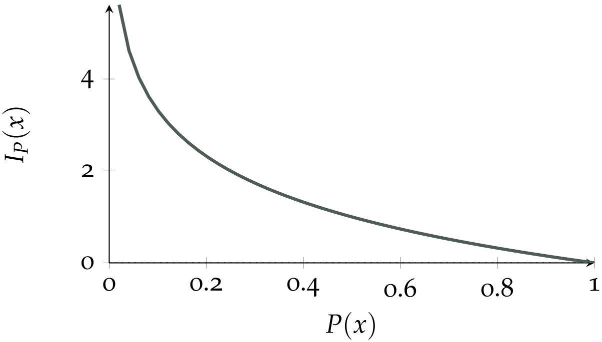
$I_p(x)$ measures how much an agent with beliefs $P \in \Delta(X)$ learns when observing $x \in X$.
Alt.: how surprised that agent would be when $x \in X$ actually happened.

- if $P(x) = 1$, then $I_P(x) = 0$
- if $P(x_1) < P(x_2)$, then $I_P(x_1) > I_P(x_2)$
- if x_1 and x_2 are independent, then $I_P(x_1 \& x_2) = I_P(x_1) + I_P(x_2)$

Information content

Let $P \in \Delta(X)$ be a probability distribution over (finite) set X. For event $x \in X$, the **information content** $I_P(x)$ of x (a.k.a. **surprisal** of x) is:

$$I_P(x) = -\log_2 P(x)$$



Measures of expected surprisal

general template: $\mathbb{E}_{P_g} I_{P_o}$

measure	P_g	P_o	paraphrase
entropy	P	P	average surprisal for true beliefs
cross entropy	P	Q	avrg. surprisal for beliefs Q & true P

Measures of expected difference in surprisal

general template: $\mathbb{E}_{P_g}(I_{P_o} - I_{P_g})$

measure	P_g	P_o	paraphrase
KL-divergence	P	Q	avrg. excess surprisal under Q and true P
mutual inform.	$R \in \Delta(X \times Y)$	$S(x, y) = R(x)R(y)$	avrg. excess surprisal for belief in independence X and Y