Worksheet: genetic diversity and half-life
Math 102 Section 102
Oct. 30, 2018

64

(**Genetic diversity**) 1. Entropy and other information theory approaches are used to quantify how genetically diverse a population is. The (Shannon) entropy represents the amount of uncertainty in the information. A given gene can be either on with probability $0 \leq x \leq 1$, or off with probability $0 \leq x \leq 1$. The Shannon entropy for the gene is

$$H = -x \log x - y \log y$$

Find x and y so that the genotype is the most unpredictable.
(*hint*: It amounts to finding when the Shannon entropy is the greatest. Make sure you understand why this is the case.)

(**Half-life and carbon dating**) 2. Researchers at Charlie Lake in BC have found some artifacts. For instance, a butchered bison bone that contains 0.25 mg of ^{14}C isotope. A comparable bone of a bison alive today contains about 1 mg of ^{14}C. We know that the half-life of ^{14}C is 5730 years. How many years ago could human habitation be dated back to in this region?

Solution

1. (Genetic diversity)

Notice a gene can only have one of two states: either on or off. So

$$x + y = 1 \quad \text{(constraint)}$$

Genotype is the most unpredictable = maximal amount of uncertainty in the information = maximize the Shannon entropy.

The question is essentially: ↙ log interpreted as ln.

$$\text{maximize} \quad H = -x\log x - y\log y$$
$$\text{under the constraint} \quad x + y = 1$$

$x + y = 1 \Rightarrow y = 1 - x$. Plug into the objective function

$$H = \underbrace{-x\log x}_{\text{Part 1}} \underbrace{-(1-x)\log(1-x)}_{\text{Part 2}}$$

chain rule: Don't forget "−" sign!

$$H'(x) = \underbrace{-\log x - x\cdot\frac{1}{x}}_{\substack{\text{Part 1} \\ \text{product rule}}} \underbrace{-(-1)\log(1-x) - (1-x)\cdot\frac{-1}{1-x}}_{\substack{\text{Part 2} \\ \text{product rule + chain rule}}}$$

$$= -\log x \cancel{-1} + \log(1-x) \cancel{+1}$$

$$= \log\frac{1-x}{x}$$

$$H'(x)=0 \Rightarrow \log\frac{1-x}{x}=0 \Rightarrow \frac{1-x}{x}=1$$

$$\Rightarrow x=\frac{1}{2} \quad \text{—— C.P.}$$

$$H''(x) = -\frac{1}{x} + \frac{-1}{1-x}$$

$$= \frac{-(1-x)-x}{x(1-x)} = \frac{-1}{x(1-x)}$$

Since $0 \le x \le 1$, $H''(x) < 0$. $H(x)$ is concave down. By SDT, H attains a local max at $x=\frac{1}{2}$. It is the only C.P. in $0<x<1$. So it's the global max.

Hence, the gene is the most unpredictable when there is an equal chance for it to be on or off:

$$x = y = \frac{1}{2}.$$

2. (Carbon dating)

Originally, the bison bone contains an amount of ^{14}C:

$$Q_{old} = 1 \text{ mg}.$$

With every $t_{1/2} = 5730$ years, the amount of ^{14}C drops by half. Thus

$$Q_{now} = Q_{old} \cdot \left(\frac{1}{2}\right)^{\frac{t}{t_{1/2}}}$$

$$\Rightarrow 0.25 \text{ mg} = 1 \text{ mg} \cdot \left(\frac{1}{2}\right)^{\frac{t}{t_{1/2}}}$$

$$\Rightarrow \quad \left(\frac{1}{2}\right)^{\frac{t}{t_{1/2}}} = \frac{1}{4}$$

$$\Rightarrow \quad \frac{t}{t_{1/2}} = 2$$

$$\Rightarrow \quad t = 2t_{1/2} = 11460 \text{ years.}$$

So the butchered bone indicates that human activity in this region can be dated back to at least 11460 years ago.

Remark: notice that all the units also work out consistently.