Shannon's Fundamental Equations

by Tom Schneider

2011 Feb 01

Here's an equation from Shannon's famous 1948 paper [1]:

$$H = -\sum_{i=1}^{M} p_i \log_2 p_i \qquad \text{bits per second} \tag{1}$$

where p_i are the probabilities of symbols in a message. Shannon used that equation for the *uncertainty* to define information as the decrease of uncertainty at the receiver from before receiving the message to after:

$$R = H_{\text{before}} - H_{\text{after}} \qquad \text{bits per second} \tag{2}$$

Then in 1949 Shannon used a beautiful geometric method to derive the channel capacity equation [2]:

$$C = W \log_2 \left(\frac{P}{N} + 1 \right) \qquad \text{bits per second} \tag{3}$$

where W is the bandwidth of the communication, P is the power received N is the noise at the receiver.

Finally, Shannon related the information R (equation 2) to the channel capacity (equation 3) in the famous channel capacity theorem. This theorem says that R cannot exceed C. But if $R \leq C$ then one may have as few errors in the communication as desired. Even Shannon found this surprising [2].

References

[1] C. E. Shannon. A Mathematical Theory of Communication. *Bell System Tech. J.*, 27:379–423, 623–656, 1948. http://tinyurl.com/Shannon1948.

[2] C. E. Shannon. Communication in the Presence of Noise. *Proc. IRE*, 37:10–21, 1949.