Shannon Entropy of Information:

. Information tells us something unexpected, i.e., there is a "surprise". The Ford Model T line held no surprise…every car coming off the line was an identical Black Model T, every flywheel magneto was the same with 100% probability, and hence no information was to be gained by looking at the next car or component coming off the line. But what if you were told that it was July 4th in Dallas and there was four feet of snow on the ground…this highly improbable event would be very surprising and hence convey huge information. Therefore we conclude that the *amount* of Information is inversely related to the probability of the event. It is also reasonable that, whatever the functional form of Information may be , if two independent events, 1 and 2 happen, the total information is the sum of their separate Information I_1 and I_2, i.e., $I_{1\&2}=I_1+I_2$. But the probability of independent events 1 and 2 both happening is the product of their probabilities $p_{1\&2}= p_1p_2$. So we need some function for Information I such that: $I_{1\&2}(p_1p_2)=I_1(p_1)+I_2(p_2)$ and the only function which satisfies this requirement is $I = \log(p)$ since $\log(p_1p_2) =\log(p_1)+\log(p_2)$. Therefore $I(p)=\log(p)$. But since we want the Information to be larger if the probability is smaller we will define $I(p)=\log(1/p) = -\log(p)$ which still satsifies $\log(1/p_1p_2)=\log(1/p_1)+\log(1/p_2)$. The *average* amount of information among λ choices is, like any other average, just the sum of the probability of each choice times the value of each choice:

$$H= -\sum_{i=1}^{\lambda} p_iI_i= -\sum_{i=1}^{\lambda} p_i\log_2 p_i$$

We can compute the number of typical sequences which the customer transmit in (1) in terms of Shannon Entropy:

The market makes λ Choices monthly (in this case, the unit of time is a month) and for simplicity we will assume the choices are either A or B, hence $\lambda = \lambda_A + \lambda_B$. The number of *distinct* sequences or "messages" sent by the market, to be satisfied by the company, is calculated by the usual combinatorial formula[i]:

Number of Distinct Messages = M

$$M=\frac{\lambda!}{\lambda_A!\lambda_B!}=\frac{\lambda!}{\lambda_A!(\lambda-\lambda_A)!}$$

We will follow Boltzmann by taking the logarithm of the number of states, which in the Microeconomic case is the number of distinct messages from the market: According to Stirling's formula, to first order[ii]:

$\log_2\lambda! \cong \lambda\log_2\lambda-\lambda$, note that $\lambda=(\lambda-\lambda_B)+\lambda_A=\lambda_B+\lambda_A$

$$\log_2 M=\left(\lambda\log_2\lambda -\lambda_A\log_2\lambda_A-(\lambda-\lambda_A)\log_2(\lambda-\lambda_A)\right)$$

$$\log_2 M=\left((\lambda-\lambda_A)+\lambda_A\right)\log_2\lambda -\lambda_A\log_2\lambda_A-$$

$$(\lambda-\lambda_A)\log_2(\lambda-\lambda_A)$$

$$\log_2 M=-\left((\lambda-\lambda_A)\log_2\left(\frac{\lambda-\lambda_A}{\lambda}\right)+\lambda_A\log_2\left(\frac{\lambda_A}{\lambda}\right)\right),$$

multiplying by $\frac{\lambda}{\lambda}$,obtain:

$$\log_2 M=-\lambda\left(\left(\frac{\lambda-\lambda_A}{\lambda}\right)\log_2\left(\frac{\lambda-\lambda_A}{\lambda}\right)+\left(\frac{\lambda_A}{\lambda}\right)\log_2\left(\frac{\lambda_A}{\lambda}\right)\right),$$

$$\text{let } p_A=\frac{\lambda_A}{\lambda}, p_B=\frac{\lambda-\lambda_A}{\lambda}$$

$$\log_2 M=\lambda\left(-\left\{p_1\log_2 p_1+p_2\log_2 p_2\right\}\right)\rightarrow$$

$$\lambda\left\{-\sum_{i=1}^{m} p_i\log_2 p_i\right\}=\lambda H_m \text{ for m products,}$$

$M=2^{\lambda H_m} \cong$ Number of Distinct Messages M due to m different products which are supplied to customers. Because of the approximation in Stirling's formula, we are missing a few improbable messages,hence there are $2^{\lambda H_m}$ typical messages which is known as the "typical set"

Notice that Shannon's equation for Information emerged naturally. The market is making λ variety choices per month, selected from one of the m products, each of the λ events per month containing information of H_m bits. The M messages per month corresponds to the number of unique states per month .

$$H_m=-\sum_{i=1}^{m} p_i\log p_i = \text{Information in Bits per Choice}$$

$$\text{Transmission Rate of Market} \rightarrow \lambda H_m = \left(\frac{\text{Choices}}{\text{Month}}\right)\left(\frac{\text{Bits}}{\text{Choice}}\right)\rightarrow \text{ Variety Bits per Month}$$

Thus the market is acting like a communication system, transmitting λH_M bits of information per month about the variety of products it wants to buy which the company presently offers. Referring to the early automotive market, initially the market demanded utility transportation and Ford responded with m =1 in the form of the Model T. As the technology of cars improved from 1908 to 1925, Ford continued on with m=1 whereas the market demanded variety as brilliantly offered by Sloan of GM where m>5, and the seemingly impregnable Model T was quickly destroyed[iii].

---

[i] Walpole,Ronald et al (2002) *Probability and Statistics for Engineers and Scientists* p.37

[ii] Stirling's formula logD!=DlogD-D is only in error by 1% when the number of products shipped per month is D=10, and of course is entirely negligible for most companies when D>>10. See Reif, F 1965, *Fundamentals of Statistical and Thermal Physics*, pp 613-614 for an investigation of the accuracy of Stirling's formula.

[iii] MIT LAI paper 2007 op cit