Shannon entropy

A series of messages $\{\mathcal{M}_n\}$ $(n = 1 \ldots N)$ will be transmitted over a telephone line. Each message occurs with probability P_n.

1. How much information, I_m is required to transmit message m?
2. What is the average information per message?
3. Determine the information content of a message if all messages are equally probable (*i.e.* $P_n = 1/N$).
4. If $P_m > 1/N$ for some m, does the information content of m increase, or decrease? Why?
5. What is the probability of a pair of messages, m and n?
6. What is the information content I_{mn} of that pair of messages?
7. Prove the identity $I_m + I_n = I_{mn}$ and give an intuitive explanation of its meaning. Does that identity determine the functional form of I_m?