MASSACHUSETTS INSTITUTE OF TECHNOLOGY
Department of Electrical Engineering and Computer Science
Department of Mechanical Engineering

6.050J/2.110J **Information and Entropy** **Spring 2003**

Problem Set 7 Solutions

Solution to Problem 1: Uncertain Employment

Solution to Problem 1, part a.

i. We will need to rearrange Equations 7–1 and 7–2. First we get the following from 7–2

$$p(F) = 1.9 - 2p(E) \tag{7–3}$$

and substituting this into 7–1 we have

$$\begin{aligned} 1 &= p(E) + p(F) + p(U) \\ &= p(E) + 1.9 - 2p(E) + p(U) \\ &= 1.9 - p(E) + p(U) \qquad (7\text{–}4) \end{aligned}$$

$$p(E) = 0.9 + p(U) \tag{7–5}$$

Once $p(E)$ and $p(U)$ are determined, $p(F)$ is determined also:

$$\begin{aligned} p(F) &= 0.1 - 2p(U) \\ &= 1.9 - 2p(E) \qquad (7\text{–}6) \end{aligned}$$

And so, since both $p(U)$, $p(F)$, and $p(E)$ must be between 0 and 1, we see that $p(E)$ must be between 0.9 and 0.95. Furthermore, $p(U)$ can only be between 0 and 0.05. These are reasonable answers: if the average employment is to be kept high month to month, then we would expect that fully-employed rate would be high and unemployment rate would be low.

ii. The equation for the entropy is as follows:

$$\begin{aligned} H &= p(E)\log_2\left(\frac{1}{p(E)}\right) + p(F)\log_2\left(\frac{1}{p(F)}\right) + p(U)\log_2\left(\frac{1}{p(U)}\right) \\ &= p(E)\log_2\left(\frac{1}{p(E)}\right) + (1.9 - 2p(E))\log_2\left(\frac{1}{1.9 - 2p(E)}\right) + \\ &\quad (p(E) - 0.9)\log_2\left(\frac{1}{p(E) - 0.9}\right) \qquad (7\text{–}7) \end{aligned}$$

Plotting this for $p(F) = 0.9$ to 0.95 we get the graph shown in Figure 7–1

iii. The maximum entropy of $H = 0.4823$ bits is at $p(E) = 0.9081$, which gives values of $p(F) = 0.0838$ and $p(U) = 0.0081$.

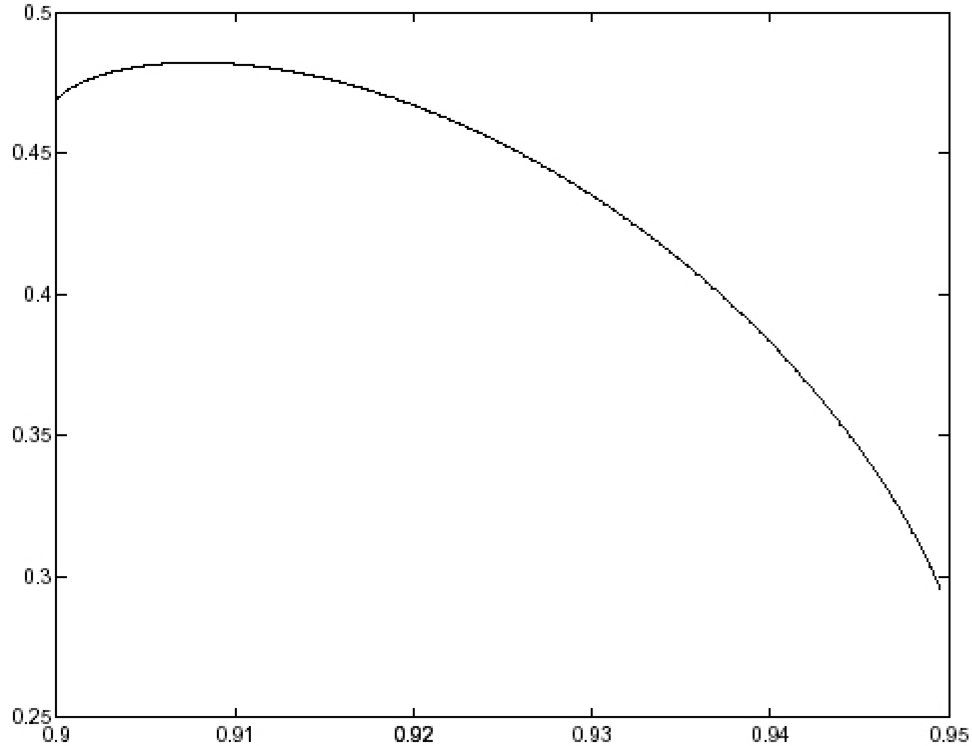

Figure 7–1: Entropy of the employment probability distribution

Solution to Problem 1, part b.

The entropy should be less than (a-iii) because, after all, that value was calculated with a procedure known as the Principle of Maximum Entropy. The maximum value of $p(E)$ consistent with the constraints is 0.95.

Solution to Problem 1, part c.

If $p(E)$ is 0.95, then $p(F) = 0$ and $p(U) = 0.05$. The entropy at this point is $H = 0.2957$ bits.

Solution to Problem 1, part d.

The constraint has changed to, counting months as i

$$0.95 = \sum_{i=1...48} \frac{p(E_i) + 0.5p(F_i)}{48} \tag{7–8}$$

(Note that a President's term has 48 months) If every citizen is unemployed in one month the constraint can still be satisfied if the employment in all the other months averages out to $(48/47) \times 0.95 = 0.97$, clearly possible if the government has the ability to adjust employment anywhere between 0 and 1. Thus the minimum employment is 0. The entropy in that month with zero employment is 0 bits.

Solution to Problem 1, part e.

If every citizen is fully employed in one month, the employment in that month is 1. The constaint can still be satisfied if the rest of the months average out to $(48 \times 0.95 - 1)/47 = 0.94$. Thus the maximum employment is 1. The entropy in this case is 0 bits.

Solution to Problem 2: Candy Raffle

Solution to Problem 2, part a.

We know that the contents of the jar have a mass of 2300 grams. The equation that governs this is

$$2.1n(T) + 1.5n(K) + 0.9n(M) = 2300 \tag{7–9}$$

The numbers compatible with this equation are calculated by dividing the weight of each piece of candy into the maximum weight of the contents:

$$\begin{aligned} n(T) &= 0 \text{ to } 1095 \\ n(K) &= 0 \text{ to } 1533 \\ n(M) &= 0 \text{ to } 2556 \end{aligned} \tag{7–10}$$

Solution to Problem 2, part b.

Knowing there are 1100 pieces of candy in the jar adds a second constraint to our system.

$$n(T) + n(K) + n(M) = 1100 \tag{7–11}$$

To calculate the percentages that maximize the uncertainty, we will cast our constraints as follows:

$$\begin{aligned} n(M) &= 1100 - n(T) - n(K) \\ 2300 &= 2.1n(T) + 1.5n(K) + 990 - 0.9n(T) - 0.9n(K) \\ 2300 &= 1.2n(T) + 0.6n(K) + 990 \end{aligned} \tag{7–12}$$

and thus, in terms of Tootsie Rolls:

$$\begin{aligned} n(K) &= 2183.33 - 2n(T) \\ n(M) &= n(T) - 1083.33 \end{aligned} \tag{7–13}$$

Since $n(K)$ cannot be less than zero, $n(T)$ cannot be greater than $2183.33/2$, or 1091 (rounding to the nearest integer). Furthermore, $n(T)$ cannot be less than 1084, because this would then make $n(M)$ less than zero. So we have bounds on $n(T)$.

$$\begin{aligned} H &= \left(\frac{n(T)}{1100}\right)\log_2\left(\frac{1100}{n(T)}\right) + \left(\frac{n(K)}{1100}\right)\log_2\left(\frac{1100}{n(K)}\right) + \left(\frac{n(M)}{1100}\right)\log_2\left(\frac{1100}{n(M)}\right) \\ H &= \left(\frac{n(T)}{1100}\right)\log_2\left(\frac{1100}{n(T)}\right) \\ &\quad + \left(\frac{2183.33 - 2n(T)}{1100}\right)\log_2\left(\frac{1100}{2183.33 - 2n(T)}\right) \\ &\quad + \left(\frac{n(T) - 1083.33}{1100}\right)\log_2\left(\frac{1100}{n(T) - 1083.33}\right) \end{aligned} \tag{7–14}$$

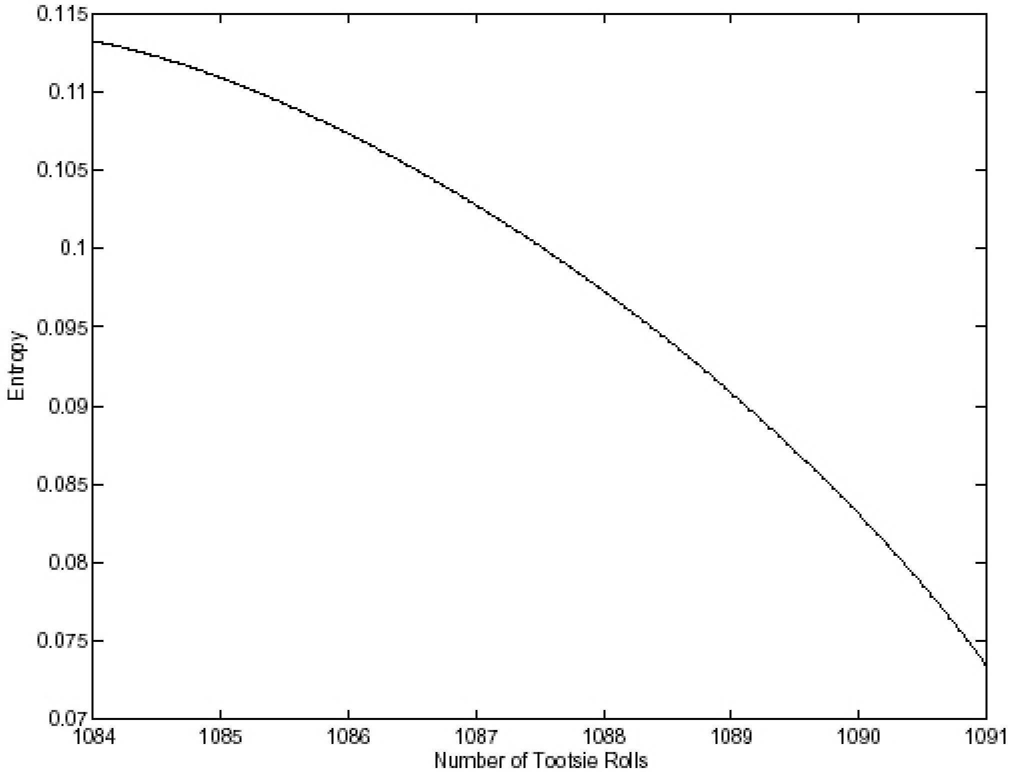

Figure 7–2: Entropy of the candy probability distribution

This graph is shown in Figure 7–2.
According to this graph, the maximum occurs at $n(T) = 1084$, with an entropy of $H = 0.1130$. This gives $n(M) = 0$, and $n(K) = 16$.

Solution to Problem 2, part c.

The percentage mass of sugar is calculated via the following equation:

$$\begin{aligned} S &= 0.85 \times 2.1 \times n(T) + 0.45 \times 0.9 \times n(M) + 0.25 \times 1.5 \times n(K) \\ &= 1.785 \times 1084 + 0.405 \times 0 + 0.375 \times 16 \\ &= 1934.94 + 0 + 6 \\ &= 1940.94 \text{ grams of sugar} \\ &= 84.3\% \text{ sugar} \end{aligned} \tag{7–15}$$

Solution to Problem 2, part d.

No, you cannot have 600 Tootsie Rolls. If you had 600 Tootsie Rolls (1260 g), would have at most 600 Hershey Kisses (900 g) for a total weight of 2160g, which does not satisfy the weight constraint of 2300g.